

SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: JONATHAN EDWARD LIGHTNER  
JOHN JOSEPH OKULEY
- (ii) TITLE OF INVENTION: GENES FOR MICROSOMAL  
FATTY ACID DELTA-12  
DESATURASES AND RELATED  
ENZYMES FROM PLANTS
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: E. I. DU PONT DE NEMOURS  
AND COMPANY
  - (B) STREET: 1007 MARKET STREET
  - (C) CITY: WILMINGTON
  - (D) STATE: DELAWARE
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 19898
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: MacIntosh
  - (C) OPERATING SYSTEM: MacIntosh 6.0
  - (D) SOFTWARE: Microsoft Word 4.0
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: U.S. 07/977,339
  - (B) FILING DATE: 17-NOV-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Siegell, Barbara C.
  - (B) REGISTRATION NUMBER: 30,684
  - (C) REFERENCE/DOCKET NUMBER: BB-1043-B
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (302) 992-4927
  - (B) TELEFAX: (302) 892-7949
  - (C) TELEX: 835420

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1372 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: p92103
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 93..1244
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGAGAGAGAG ATTCTGCGGA GGAGCTTCTT CTTCGTAGGG TGTTTCATCGT TATTAACGTT	60
ATCGCCCCTA CGTCAGCTCC ATCTCCAGAA AC ATG GGT GCA GGT GGA AGA ATG	113
Met Gly Ala Gly Gly Arg Met	
1 5	
CCG GTT CCT ACT TCT TCC AAG AAA TCG GAA ACC GAC ACC ACA AAG CGT	161
Pro Val Pro Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr Thr Lys Arg	
10 15 20	
GTG CCG TGC GAG AAA CCG CCT TTC TCG GTG GGA GAT CTG AAG AAA GCA	209
Val Pro Cys Glu Lys Pro Pro Phe Ser Val Gly Asp Leu Lys Lys Ala	
25 30 35	
ATC CCG CCG CAT TGT TTC AAA CGC TCA ATC CCT CGC TCT TTC TCC TAC	257
Ile Pro Pro His Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ser Tyr	
40 45 50 55	
CTT ATC AGT GAC ATC ATT ATA GCC TCA TGC TTC TAC TAC GTC GCC ACC	305
Leu Ile Ser Asp Ile Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala Thr	
60 65 70	
AAT TAC TTC TCT CTC CTC CCT CAG CCT CTC TCT TAC TTG GCT TGG CCA	353
Asn Tyr Phe Ser Leu Leu Pro Gln Pro Leu Ser Tyr Leu Ala Trp Pro	
75 80 85	
CTC TAT TGG GCC TGT CAA GGC TGT GTC CTA ACT GGT ATC TGG GTC ATA	401
Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Thr Gly Ile Trp Val Ile	
90 95 100	

GCC CAC GAA TGC GGT CAC CAC GCA TTC AGC GAC TAC CAA TGG CTG GAT Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu Asp 105 110 115	449
GAC ACA GTT GGT CTT ATC TTC CAT TCC TTC CTC CTC GTC CCT TAC TTC Asp Thr Val Gly Leu Ile Phe His Ser Phe Leu Leu Val Pro Tyr Phe 120 125 130 135	497
TCC TGG AAG TAT AGT CAT CGC CGT CAC CAT TCC AAC ACT GGA TCC CTC Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu 140 145 150	545
GAA AGA GAT GAA GTA TTT GTC CCA AAG CAG AAA TCA GCA ATC AAG TGG Glu Arg Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Ile Lys Trp 155 160 165	593
TAC GGG AAA TAC CTC AAC AAC CCT CTT GGA CGC ATC ATG ATG TTA ACC Tyr Gly Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met Met Leu Thr 170 175 180	641
GTC CAG TTT GTC CTC GGG TGG CCC TTG TAC TTA GCC TTT AAC GTC TCT Val Gln Phe Val Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser 185 190 195	689
GGC AGA CCG TAT GAC GGG TTC GCT TGC CAT TTC TTC CCC AAC GCT CCC Gly Arg Pro Tyr Asp Gly Phe Ala Cys His Phe Phe Pro Asn Ala Pro 200 205 210 215	737
ATC TAC AAT GAC CGA GAA CGC CTC CAG ATA TAC CTC TCT GAT GCG GGT Ile Tyr Asn Asp Arg Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala Gly 220 225 230	785
ATT CTA GCC GTC TGT TTT GGT CTT TAC CGT TAC GCT GCT GCA CAA GGG Ile Leu Ala Val Cys Phe Gly Leu Tyr Arg Tyr Ala Ala Ala Gln Gly 235 240 245	833
ATG GCC TCG ATG ATC TGC CTC TAC GGA GTA CCG CTT CTG ATA GTG AAT Met Ala Ser Met Ile Cys Leu Tyr Gly Val Pro Leu Leu Ile Val Asn 250 255 260	881
GCG TTC CTC GTC TTG ATC ACT TAC TTG CAG CAC ACT CAT CCC TCG TTG Ala Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu 265 270 275	929
CCT CAC TAC GAT TCA TCA GAG TGG GAC TGG CTC AGG GGA GCT TTG GCT Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala 280 285 290 295	977
ACC GTA GAC AGA GAC TAC GGA ATC TTG AAC AAG GTG TTC CAC AAC ATT Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile 300 305 310	1025
ACA GAC ACA CAC GTG GCT CAT CAC CTG TTC TCG ACA ATG CCG CAT TAT Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr 315 320 325	1073

AAC GCA ATG GAA GCT ACA AAG GCG ATA AAG CCA ATT CTG GGA GAC TAT 1121  
 Asn Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Asp Tyr  
 330 335 340  
  
 TAC CAG TTC GAT GGA ACA CCG TGG TAT GTA GCG ATG TAT AGG GAG GCA 1169  
 Tyr Gln Phe Asp Gly Thr Pro Trp Tyr Val Ala Met Tyr Arg Glu Ala  
 345 350 355  
  
 AAG GAG TGT ATC TAT GTA GAA CCG GAC AGG GAA GGT GAC AAG AAA GGT 1217  
 Lys Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu Gly Asp Lys Lys Gly  
 360 365 370 375  
  
 GTG TAC TGG TAC AAC AAT AAG TTA TGAGCATGAT GGTGAAGAAA TTGTCGACCT 1271  
 Val Tyr Trp Tyr Asn Asn Lys Leu  
 380  
  
 TTCTCTTGTC TGTTTGCTTT TTGTTAAAGA AGCTATGCTT CGTTTTAATA ATCTTATTGT 1331  
  
 CCATTTTGTGTT GTGTTATGAC ATTTTGGCTG CTCATTATGT T 1372

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser  
 1 5 10 15  
 Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser  
 20 25 30  
 Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser  
 35 40 45  
 Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ala Ser  
 50 55 60  
 Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro  
 65 70 75 80  
 Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val  
 85 90 95  
 Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe  
 100 105 110  
 Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser  
 115 120 125  
 Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His  
 130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys  
 145 150 155 160  
 Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu  
 165 170 175  
 Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu  
 180 185 190  
 Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys  
 195 200 205  
 His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln  
 210 215 220  
 Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr  
 225 230 235 240  
 Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly  
 245 250 255  
 Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu  
 260 265 270  
 Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp  
 275 280 285  
 Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu  
 290 295 300  
 Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu  
 305 310 315 320  
 Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile  
 325 330 335  
 Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr  
 340 345 350  
 Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp  
 355 360 365  
 Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu  
 370 375 380

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1394 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Brassica napus

(vii) IMMEDIATE SOURCE:

(B) CLONE: pCF2-165D

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 99..1250

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAGAGGAGAC AGAGACAGAG AGAGAGTTGA GAGAGCTCTC GTAGGTTATC GTATTAACGT	60
AATCTTCAAT CCCCCCTACG TCAGCCAGCT CAAGAAAC ATG GGT GCA GGT GGA	113
Met Gly Ala Gly Gly	
1 5	
AGA ATG CAA GTG TCT CCT CCC TCC AAA AAG TCT GAA ACC GAC AAC ATC	161
Arg Met Gln Val Ser Pro Pro Ser Lys Lys Ser Glu Thr Asp Asn Ile	
10 15 20	
AAG CGC GTA CCC TGC GAG ACA CCG CCC TTC ACT GTC GGA GAA CTC AAG	209
Lys Arg Val Pro Cys Glu Thr Pro Pro Phe Thr Val Gly Glu Leu Lys	
25 30 35	
AAA GCA ATC CCA CCG CAC TGT TTC AAG CGC TCG ATC CCT CGC TCT TTC	257
Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe	
40 45 50	
TCC CAC CTC ATC TGG GAC ATC ATC ATA GCC TCC TGC TTC TAC TAC GTC	305
Ser His Leu Ile Trp Asp Ile Ile Ala Ser Cys Phe Tyr Tyr Val	
55 60 65	
GCC ACC ACT TAC TTC CCT CTC CTC CCT AAC CCT CTC TCC TAC TTC GCC	353
Ala Thr Thr Tyr Phe Pro Leu Leu Pro Asn Pro Leu Ser Tyr Phe Ala	
70 75 80 85	
TGG CCT CTC TAC TGG GCC TGC CAG GGC TGC GTC CTA ACC GGC GTC TGG	401
Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Thr Gly Val Trp	
90 95 100	
GTC ATA GGC CAC GAG TGC GGC CAC GCA GGC TTC AGC GAC TAC CAG TGG	449
Val Ile Ala His Glu Cys Gly His Ala Ala Phe Ser Asp Tyr Gln Trp	
105 110 115	
CTG GAC GAC ACC GTC GGC CTC ATC TTC CAC TCC TTC CTC CTC GTC CCT	497
Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser Phe Leu Leu Val Pro	
120 125 130	

TAC TTC TCC TGG AAG TAC AGT CAT CGA CGC CAC CAT TCC AAC ACT GGC Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly 135 140 145	545
TCC CTC GAG AGA GAC GAA GTG TTT GTC CCA AGA AGA AGT CAG ACA TCA Ser Leu Glu Arg Asp Glu Val Phe Val Pro Arg Arg Ser Gln Thr Ser 150 155 160 165	593
AGT GGT ACG GCA AGT ACC TCA ACA ACC TTT GGA CGC ACC GTG ATG TTA Ser Gly Thr Ala Ser Thr Ser Thr Thr Phe Gly Arg Thr Val Met Leu 170 175 180	641
ACG GTT CAG TTC ACT CTC GGC TGG CCT TTG TAC TTA GCC TTC AAC GTC Thr Val Gln Phe Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val 185 190 195	689
TCG GGG AGA CCT TAC GAC GGC GGC TTC GCT TGC CAT TTC CAC CCC AAC Ser Gly Arg Pro Tyr Asp Gly Gly Phe Ala Cys His Phe His Pro Asn 200 205 210	737
GCT CCC ATC TAC AAC GAC CGT GAG CGT CTC CAG ATA TAC ATC TCC GAC Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp 215 220 225	785
GCT GGC ATC CTC GCC GTC TGC TAC GGT CTG CTA CCG TAC GCT GCT GTC Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu Leu Pro Tyr Ala Ala Val 230 235 240 245	833
CAA GGA GTT GCC TCG ATG GTC TGC TTC CTA CGA GTT CCT CTT CTG ATT Gln Gly Val Ala Ser Met Val Cys Phe Leu Arg Val Pro Leu Leu Ile 250 255 260	881
GTC AAC GGG TTC TTA GTT TTG ATC ACT TAC TTG CAG CAC ACG CAT CCT Val Asn Gly Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro 265 270 275	929
TCC CTG CCT CAC TAT GAC TCG TCT GAG TGG GAT TGG TTG AGG GGA GCT Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala 280 285 290	977
TTG GCC ACC GTT GAC AGA GAC TAC GGA ATC TTG AAC CAA GGC TTC CAC Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Gln Gly Phe His 295 300 305	1025
AAT ATC ACG GAC ACG CAC GAG GCG CAT CAC CTG TTC TCG ACC ATG CCG Asn Ile Thr Asp Thr His Glu Ala His His Leu Phe Ser Thr Met Pro 310 315 320 325	1073
CAT TAT CAT GCG ATG GAA GCT ACG AAG GCG ATA AAG CCG ATA CTG GGA His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly 330 335 340	1121
GAG TAT TAT CAG TTC GAT GGG ACG CCG GTG GTT AAG GCG ATG TGG AGG Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Val Val Lys Ala Met Trp Arg 345 350 355	1169

GAG GCG AAG GAG TGT ATC TAT GTG GAA CCG GAC AGG CAA GGT GAG AAG	1217
Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp Arg Gln Gly Glu Lys	
360 365 370	
AAA GGT GTG TTC TGG TAC AAC AAT AAG TTA TGAAGCAAAG AAGAACTGA	1267
Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu	
375 380	
ACCTTTCTCT TCTATCAATT GTCTTTGTTT AAGAAGCTAT GTTTCTGTTT CAATAATCTT	1327
AATTATCCAT TTTGTTGTGT TTTCTGACAT TTTGGCTAAA ATTATGTGAT GTTGGAAGTT	1387
AGTGTCT	1394

(2) INFORMATION FOR SEO ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met 1	Gly	Ala	Gly	Gly 5	Arg	Met	Gln	Val	Ser 10	Pro	Pro	Ser	Lys	Lys 15	Ser
Glu	Thr	Asp	Asn 20	Ile	Lys	Arg	Val	Pro 25	Cys	Glu	Thr	Pro	Pro 30	Phe	Thr
Val	Gly	Glu	Leu 35	Lys	Lys	Ala	Ile 40	Pro	Pro	His	Cys	Phe 45	Lys	Arg	Ser
Ile	Pro 50	Arg	Ser	Phe	Ser	His 55	Leu	Ile	Trp	Asp	Ile 60	Ile	Ile	Ala	Ser
Cys 65	Phe	Tyr	Tyr	Val	Ala 70	Thr	Thr	Tyr	Phe	Pro 75	Leu	Leu	Pro	Asn	Pro 80
Leu	Ser	Tyr	Phe	Ala 85	Trp	Pro	Leu	Tyr	Trp 90	Ala	Cys	Gln	Gly	Cys 95	Val
Leu	Thr	Gly	Val 100	Trp	Val	Ile	Ala	His 105	Glu	Cys	Gly	His	Ala 110	Ala	Phe
Ser	Asp	Tyr 115	Gln	Trp	Leu	Asp	Asp 120	Thr	Val	Gly	Leu	Ile 125	Phe	His	Ser
Phe 130	Leu	Leu	Val	Pro	Tyr	Phe 135	Ser	Trp	Lys	Tyr	Ser 140	His	Arg	Arg	His
His 145	Ser	Asn	Thr	Gly	Ser 150	Leu	Glu	Arg	Asp	Glu 155	Val	Phe	Val	Pro	Arg 160
Arg	Ser	Gln	Thr	Ser 165	Ser	Gly	Thr	Ala	Ser 170	Thr	Ser	Thr	Thr	Phe 175	Gly



Arg	Thr	Val	Met	Leu	Thr	Val	Gln	Phe	Thr	Leu	Gly	Trp	Pro	Leu	Tyr
			180					185					190		
Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly	Gly	Phe	Ala	Cys
		195					200					205			
His	Phe	His	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg	Glu	Arg	Leu	Gln
	210					215					220				
Ile	Tyr	Ile	Ser	Asp	Ala	Gly	Ile	Leu	Ala	Val	Cys	Tyr	Gly	Leu	Leu
225					230					235					240
Pro	Tyr	Ala	Ala	Val	Gln	Gly	Val	Ala	Ser	Met	Val	Cys	Phe	Leu	Arg
				245					250					255	
Val	Pro	Leu	Leu	Ile	Val	Asn	Gly	Phe	Leu	Val	Leu	Ile	Thr	Tyr	Leu
			260					265					270		
Gln	His	Thr	His	Pro	Ser	Leu	Pro	His	Tyr	Asp	Ser	Ser	Glu	Trp	Asp
		275					280					285			
Trp	Leu	Arg	Gly	Ala	Leu	Ala	Thr	Val	Asp	Arg	Asp	Tyr	Gly	Ile	Leu
	290					295					300				
Asn	Gln	Gly	Phe	His	Asn	Ile	Thr	Asp	Thr	His	Glu	Ala	His	His	Leu
305					310					315					320
Phe	Ser	Thr	Met	Pro	His	Tyr	His	Ala	Met	Glu	Ala	Thr	Lys	Ala	Ile
				325					330					335	
Lys	Pro	Ile	Leu	Gly	Glu	Tyr	Tyr	Gln	Phe	Asp	Gly	Thr	Pro	Val	Val
			340					345					350		
Lys	Ala	Met	Trp	Arg	Glu	Ala	Lys	Glu	Cys	Ile	Tyr	Val	Glu	Pro	Asp
		355					360					365			
Arg	Gln	Gly	Glu	Lys	Lys	Gly	Val	Phe	Trp	Tyr	Asn	Asn	Lys	Leu	
	370					375					380				

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1462 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: pSF2-165K

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 108..1247

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATATACTA ATATTTGCTT GTATTGATAG CCCCTCCGTT CCCAAGAGTA TAAAACTGCA	60
TCGAATAATA CAAGCCACTA GGCATGGGTC TAGCAAAGGA AACAAACA ATG GGA GGT	116
Met Gly Gly	1
AGA GGT CGT GTG GCC AAA GTG GAA GTT CAA GGG AAG AAG CCT CTC TCA	164
Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys Pro Leu Ser	5 10 15
AGG GTT CCA AAC ACA AAG CCA CCA TTC ACT GTT GGC CAA CTC AAG AAA	212
Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys	20 25 30 35
GCA ATT CCA CCA CAC TGC TTT CAG CGC TCC CTC CTC ACT TCA TTC TCC	260
Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr Ser Phe Ser	40 45 50
TAT GTT GTT TAT GAC CTT TCA TTT GCC TTC ATT TTC TAC ATT GCC ACC	308
Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr Ile Ala Thr	55 60 65
ACC TAC TTC CAC CTC CTT CCT CAA CCC TTT TCC CTC ATT GCA TGG CCA	356
Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile Ala Trp Pro	70 75 80
ATC TAT TGG GTT CTC CAA GGT TGC CTT CTC ACT GGT GTG TGG GTG ATT	404
Ile Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr Gly Val Trp Val Ile	85 90 95
GCT CAC GAG TGT GGT CAC CAT GCC TTC AGC AAG TAC CAA TGG GTT GAT	452
Ala His Glu Cys Gly His His Ala Phe Ser Lys Tyr Gln Trp Val Asp	100 105 110 115
GAT GTT GTG GGT TTG ACC CTT CAC TCA ACA CTT TTA GTC CCT TAT TTC	500
Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val Pro Tyr Phe	120 125 130
TCA TGG AAA ATA AGC CAT CGC CGC CAT CAC TCC AAC ACA GGT TCC CTT	548
Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu	135 140 145
GAC CGT GAT GAA GTG TTT GTC CCA AAA CCA AAA TCC AAA GTT GCA TGG	596
Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp	150 155 160

TTT TCC AAG TAC TTA AAC AAC CCT CTA GGA AGG GCT GTT TCT CTT CTC	644
Phe Ser Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val Ser Leu Leu	
165 170 175	
GTC ACA CTC ACA ATA GGG TGG CCT ATG TAT TTA GCC TTC AAT GTC TCT	692
Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser	
180 185 190 195	
GGT AGA CCC TAT GAT AGT TTT GCA AGC CAC TAC CAC CCT TAT GCT CCC	740
Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro Tyr Ala Pro	
200 205 210	
ATA TAT TCT AAC CGT GAG AGG CTT CTG ATC TAT GTC TCT GAT GTT GCT	788
Ile Tyr Ser Asn Arg Glu Arg Leu Leu Ile Tyr Val Ser Asp Val Ala	
215 220 225	
TTG TTT TCT GTG ACT TAC TCT CTC TAC CGT GTT GCA ACC CTG AAA GGG	836
Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr Leu Lys Gly	
230 235 240	
TTG GTT TGG CTG CTA TGT GTT TAT GGG GTG CCT TTG CTC ATT GTG AAC	884
Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn	
245 250 255	
GGT TTT CTT GTG ACT ATC ACA TAT TTG CAG CAC ACA CAC TTT GCC TTG	932
Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu	
260 265 270 275	
CCT CAT TAC GAT TCA TCA GAA TGG GAC TGG CTG AAG GGA GCT TTG GCA	980
Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala	
280 285 290	
ACT ATG GAC AGA GAT TAT GGG ATT CTG AAC AAG GTG TTT CAT CAC ATA	1028
Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His His Ile	
295 300 305	
ACT GAT ACT CAT GTG GCT CAC CAT CTC TTC TCT ACA ATG CCA CAT TAC	1076
Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr	
310 315 320	
CAT GCA ATG GAG GCA ACC AAT GCA ATC AAG CCA ATA TTG GGT GAG TAC	1124
His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr	
325 330 335	
TAC CAA TTT GAT GAC ACA CCA TTT TAC AAG GCA CTG TGG AGA GAA GCG	1172
Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala	
340 345 350 355	
AGA GAG TGC CTC TAT GTG GAG CCA GAT GAA GGA ACA TCC GAG AAG GGC	1220
Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly	
360 365 370	
GTG TAT TGG TAC AGG AAC AAG TAT TGATGGAGCA ACCAATGGGC CATAGTGGGA	1274
Val Tyr Trp Tyr Arg Asn Lys Tyr	
375 380	
GTTATGGAAG TTTTGTGTCATG TATTAGTACA TAATTAGTAG AATGTTATAA ATAAGTGGAT	1334

TTGCCGCGTA ATGACTTTGT GTGTATTGTG AAACAGCTTG TTGCGATCAT GGTATAATG 1394  
TAAAAATAAT TCTGGTATTA ATTACATGTG GAAAGTGTTT TGCTTATAGC TTTCTGCCTA 1454  
AAAAAAA 1462

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 379 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Gly Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys  
1 5 10 15  
Pro Leu Ser Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln  
20 25 30  
Leu Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr  
35 40 45  
Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr  
50 55 60  
Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile  
65 70 75 80  
Ala Trp Pro Ile Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr Gly Val  
85 90 95  
Trp Val Ile Ala His Glu Cys Gly His His Ala Phe Ser Lys Tyr Gln  
100 105 110  
Trp Val Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val  
115 120 125  
Pro Tyr Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr  
130 135 140  
Gly Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys  
145 150 155 160  
Val Ala Trp Phe Ser Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val  
165 170 175  
Ser Leu Leu Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe  
180 185 190  
Asn Val Ser Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro  
195 200 205

Tyr Ala Pro Ile Tyr Ser Asn Arg Glu Arg Leu Leu Ile Tyr Val Ser  
 210 215 220  
 Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr  
 225 230 235 240  
 Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu  
 245 250 255  
 Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His  
 260 265 270  
 Phe Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly  
 275 280 285  
 Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe  
 290 295 300  
 His His Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met  
 305 310 315 320  
 Pro His Tyr His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu  
 325 330 335  
 Gly Glu Tyr Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp  
 340 345 350  
 Arg Glu Ala Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser  
 355 360 365  
 Glu Lys Gly Val Tyr Trp Tyr Arg Asn Lys Tyr  
 370 375

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1790 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Zea mays
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: pFad2#1
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 165..1328

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGCCTCTCC CCTCCCTCCT CCCTGCAAAT CCTGCAGACA CCACCGCTCG TTTTCTCTC	60
CGGGACAGGA GAAAAGGGGA GAGAGAGGTG AGGCGCGGTG TCCGCCCCGAT CTGCTCTGCC	120
CCGACGCAGC TGTTACGACC TCCTCAGTCT CAGTCAGGAG CAAG ATG GGT GCC GGC	176
Met Gly Ala Gly	
1	
GGC AGG ATG ACC GAG AAG GAG CGG GAG AAG CAG GAG CAG CTC GCC CGA	224
Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu Gln Leu Ala Arg	
5 10 15 20	
GCT ACC GGT GGC GCC GCG ATG CAG CGG TCG CCG GTG GAG AAG CCT CCG	272
Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val Glu Lys Pro Pro	
25 30 35	
TTC ACT CTG GGT CAG ATC AAG AAG GCC ATC CCG CCA CAC TGC TTC GAG	320
Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro His Cys Phe Glu	
40 45 50	
CGC TCG GTG CTC AAG TCC TTC TCG TAC GTG GTC CAC GAC CTG GTG ATC	368
Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile	
55 60 65	
GCC GCG GCG CTC CTC TAC TTC GCG CTG GCC ATC ATA CCG GCG CTC CCA	416
Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile Pro Ala Leu Pro	
70 75 80	
AGC CCG CTC CGC TAC GCC GCC TGG CCG CTG TAC TGG ATC GCG CAG GGG	464
Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp Ile Ala Gln Gly	
85 90 95 100	
TGC GTG TGC ACC GGC GTG TGG GTC ATC GCG CAC GAG TGC GGC CAC CAC	512
Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His	
105 110 115	
GCC TTC TCG GAC TAC TCG CTC CTG GAC GAC GTG GTC GGC CTG GTG CTG	560
Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Val Val Gly Leu Val Leu	
120 125 130	
CAC TCG TCG CTC ATG GTG CCC TAC TTC TCG TGG AAG TAC AGC CAC CGG	608
His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg	
135 140 145	
CGC CAC CAC TCC AAC ACG GGG TCC CTG GAG CGC GAC GAG GTG TTC GTG	656
Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val	
150 155 160	
CCC AAG AAG AAG GAG GCG CTG CCG TGG TAC ACC CCG TAC GTG TAC AAC	704
Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro Tyr Val Tyr Asn	
165 170 175 180	

AAC CCG GTC GGC CGG GTG GTG CAC ATC GTG GTG CAG CTC ACC CTC GGG Asn Pro Val Gly Arg Val Val His Ile Val Val Gln Leu Thr Leu Gly 185 190 195	752
TGG CCG CTG TAC CTG GCG ACC AAC GCG TCG GGG CGG CCG TAC CCG CGC Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg Pro Tyr Pro Arg 200 205 210	800
TTC GCC TGC CAC TTC GAC CCC TAC GGC CCC ATC TAC AAC GAC CGG GAG Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu 215 220 225	848
CGC GCC CAG ATC TTC GTC TCG GAC GCC GGC GTC GTG GCC GTG GCG TTC Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val Ala Val Ala Phe 230 235 240	896
GGG CTG TAC AAG CTG GCG GCG GCG TTC GGG GTC TGG TGG GTG GTG CGC Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp Trp Val Val Arg 245 250 255 260	944
GTG TAC GCC GTG CCG CTG CTG ATC GTG AAC GCG TGG CTG GTG CTC ATC Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile 265 270 275	992
ACC TAC CTG CAG CAC ACC CAC CCG TCG CTC CCC CAC TAC GAC TCG AGC Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser 280 285 290	1040
GAG TGG GAC TGG CTG CGC GGC GCG CTG GCC ACC ATG GAC CGC GAC TAC Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr 295 300 305	1088
GGC ATC CTC AAC CGC GTG TTC CAC AAC ATC ACG GAC ACG CAC GTC GCG Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp Thr His Val Ala 310 315 320	1136
CAC CAC CTC TTC TCC ACC ATG CCG CAC TAC CAC GCC ATG GAG GCC ACC His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr 325 330 335 340	1184
AAG GCG ATC AGG CCC ATC CTC GGC GAC TAC TAC CAC TTC GAC CCG ACC Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Pro Thr 345 350 355	1232
CCT GTC GCC AAG GCG ACC TGG CGC GAG GCC GGG GAA TGC ATC TAC GTC Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu Cys Ile Tyr Val 360 365 370	1280
GAG CCC GAG GAC CGC AAG GGC GTC TTC TGG TAC AAC AAG AAG TTC TAGCCGCCGC Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn Lys Lys Phe 375 380 385	1335
CGCTCGCAGA GCTGAGGACG CTACCGTAGG AATGGGAGCA GAAACCAGGA GGAGGAGACG	1395
GTACTCGCCC CAAAGTCTCC GTCAACCTAT CTAATCGTTA GTCGTCAGTC TTTTAGACGG	1455
GAAGAGAGAT CATTTGGGCA CAGAGACGAA GGCTTACTGC AGTGCCATCG CTAGAGCTGC	1515

CATCAAGTAC AAGTAGGCAA ATTCGTCAAC TTAGTGTGTC CCATGTTGTT TTTCTTAGTC 1575  
 GTCCGCTGCT GTAGGCTTTC CGGCGGCGGT CGTTTGTGTG GTTGGCATCC GTGGCCATGC 1635  
 CTGTGCGTGC GTGGCCGCGC TTGTCGTGTG CGTCTGTCGT CGCGTTGGCG TCGTCTCTTC 1695  
 GTGCTCCCCG TGTGTTGTTG TAAAACAAGA AGATGTTTTTC TGGTGTCTTT GCGGAATAA 1755  
 CAGATCGTCC GAACGAAAAA AAAAAAAAAA AAAA 1790

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu  
 1 5 10 15  
 Gln Leu Ala Arg Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val  
 20 25 30  
 Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro  
 35 40 45  
 His Cys Phe Glu Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His  
 50 55 60  
 Asp Leu Val Ile Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile  
 65 70 75 80  
 Pro Ala Leu Pro Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp  
 85 90 95  
 Ile Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu  
 100 105 110  
 Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Val Val  
 115 120 125  
 Gly Leu Val Leu His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys  
 130 135 140  
 Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp  
 145 150 155 160  
 Glu Val Phe Val Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro  
 165 170 175  
 Tyr Val Tyr Asn Asn Pro Val Gly Arg Val Val His Ile Val Val Gln  
 180 185 190



Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg  
 195 200 205  
 Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr  
 210 215 220  
 Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val  
 225 230 235 240  
 Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp  
 245 250 255  
 Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp  
 260 265 270  
 Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His  
 275 280 285  
 Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met  
 290 295 300  
 Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp  
 305 310 315 320  
 Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala  
 325 330 335  
 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His  
 340 345 350  
 Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu  
 355 360 365  
 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn  
 370 375 380  
 Lys Lys Phe  
 385

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 673 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Ricinus communis

(vii) IMMEDIATE SOURCE:  
(B) CLONE: pRF2-1C

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..673

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGG GTG ATG GCG CAT GAT TGT GGG CAC CAT GCC TTC AGT GAC TAT CAA	48
Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln	
1 5 10 15	
TTG CTT GAT GAT GTA GTT GGT CTT ATC CTA CAC TCC TGT CTC CTT GTC	96
Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val	
20 25 30	
CCT TAT TTT TCA TGG AAA CAC AGC CAT CGC CGA CAT CAT TCC AAC ACA	144
Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr	
35 40 45	
GGG TCC CTG GAA CGG GAT GAA GTG TTT GTT CCC AAG AAG AAA TCT AGT	192
Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser	
50 55 60	
ATC CGT TGG TAT TCC AAA TAC CTC AAC AAC CCT CCA GGT CGT ATC ATG	240
Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met	
65 70 75 80	
ACA ATT GCC GTC ACA CTT TCA CTT GGC TGG CCT CTG TAC CTA GCA TTC	288
Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe	
85 90 95	
AAT GTT TCA GGC AGG CCA TAT GAT CGG TTC GCC TGC CAC TAT GAC CCA	336
Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro	
100 105 110	
TAT GGC CCG ATC TAC AAT GAT CGC GAG CGA ATC GAG ATA TTC ATA TCA	384
Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser	
115 120 125	
GAT GCT GGT GTT CTT GCT GTC ACT TTT GGT CTC TAC CAA CTT GCT ATA	432
Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile	
130 135 140	
GCG AAG GGG CTT GCT TGG GTT GTC TGT GTA TAT GGA GTG CCA TTG TTG	480
Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu	
145 150 155 160	
GTG GTG AAT TCA TTC CTT GTT CTG ATC ACA TTT CTG CAG CAT ACT CAC	528
Val Val Asn Ser Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His	
165 170 175	
CCT GCA TTG CCA CAT TAT GAT TCG TCG GAG TGG GAC TGG CTA AGA GGA	576
Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly	
180 185 190	



(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 224 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp 1	Val	Met	Ala	His 5	Asp	Cys	Gly	His	His 10	Ala	Phe	Ser	Asp	Tyr 15	Gln
Leu	Leu	Asp	Asp 20	Val	Val	Gly	Leu	Ile 25	Leu	His	Ser	Cys	Leu 30	Leu	Val
Pro	Tyr	Phe 35	Ser	Trp	Lys	His	Ser 40	His	Arg	Arg	His	His 45	Ser	Asn	Thr
Gly	Ser 50	Leu	Glu	Arg	Asp	Glu 55	Val	Phe	Val	Pro	Lys 60	Lys	Lys	Ser	Ser
Ile 65	Arg	Trp	Tyr	Ser	Lys 70	Tyr	Leu	Asn	Asn	Pro 75	Pro	Gly	Arg	Ile	Met 80
Thr	Ile	Ala	Val	Thr 85	Leu	Ser	Leu	Gly	Trp 90	Pro	Leu	Tyr	Leu	Ala 95	Phe
Asn	Val	Ser	Gly 100	Arg	Pro	Tyr	Asp	Arg 105	Phe	Ala	Cys	His	Tyr 110	Asp	Pro
Tyr	Gly	Pro 115	Ile	Tyr	Asn	Asp	Arg 120	Glu	Arg	Ile	Glu	Ile 125	Phe	Ile	Ser
Asp 130	Ala	Gly	Val	Leu	Ala	Val 135	Thr	Phe	Gly	Leu	Tyr 140	Gln	Leu	Ala	Ile
Ala 145	Lys	Gly	Leu	Ala	Trp 150	Val	Val	Cys	Val	Tyr 155	Gly	Val	Pro	Leu	Leu 160
Val	Val	Asn	Ser	Phe 165	Leu	Val	Leu	Ile	Thr 170	Phe	Leu	Gln	His	Thr 175	His
Pro	Ala	Leu 180	Pro	His	Tyr	Asp	Ser	Ser 185	Glu	Trp	Asp	Trp	Leu 190	Arg	Gly
Ala	Leu	Ala 195	Thr	Val	Asp	Arg	Asp 200	Tyr	Gly	Ile	Leu	Asn 205	Lys	Val	Phe

His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro  
 210 215 220

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1369 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Ricinus communis

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: pRF197c-42

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 184..1347

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGGCCGGGAT TCCGGTTTTT ACACAAATTT GCAAAAAATG CATGATTTCA CCTCAAATCA	60
AACACCACAC CTTATAACTT AGTCTTAAGA GAGAGAGAGA GAGGAGACAT TTCTCTTCTC	120
TGAGATGAGC ACTTCTCTTC CAGACATCGA AGCCTCAGGA AAGTGCTTGA GAAGAGCTTG	180
AGA ATG GGA GGT GGT GGT CGC ATG TCT ACT GTC ATA ATC AGC AAC AAC	228
Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn	
1 5 10 15	
AGT GAG AAG AAA GGA GGA AGC AGC CAC CTG GAG CGA GCG CCG CAC ACG	276
Ser Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr	
20 25 30	
AAG CCT CCT TAC ACA CTT GGT AAC CTC AAG AGA GCC ATC CCA CCC CAT	324
Lys Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His	
35 40 45	
TGC TTT GAA CGC TCT TTT GTG CGC TCA TTC TCC AAT TTT GCC TAT AAT	372
Cys Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Asn Phe Ala Tyr Asn	
50 55 60	
TTC TGC TTA AGT TTT CTT TCC TAC TCG ATC GCC ACC AAC TTC TTC CCT	420
Phe Cys Leu Ser Phe Leu Ser Tyr Ser Ile Ala Thr Asn Phe Phe Pro	
65 70 75	

TAC	ATC	TCT	TCT	CCG	CTC	TCG	TAT	GTC	GCT	TGG	CTG	GTT	TAC	TGG	CTC	468
Tyr	Ile	Ser	Ser	Pro	Leu	Ser	Tyr	Val	Ala	Trp	Leu	Val	Tyr	Trp	Leu	
80					85					90					95	
TTC	CAA	GGC	TGC	ATT	CTC	ACT	GGT	CTT	TGG	GTC	ATC	GGC	CAT	GAA	TGT	516
Phe	Gln	Gly	Cys	Ile	Leu	Thr	Gly	Leu	Trp	Val	Ile	Gly	His	Glu	Cys	
				100					105					110		
GGC	CAT	CAT	GCT	TTT	AGT	GAG	TAT	CAG	CTG	GCT	GAT	GAC	ATT	GTT	GGC	564
Gly	His	His	Ala	Phe	Ser	Glu	Tyr	Gln	Leu	Ala	Asp	Asp	Ile	Val	Gly	
			115					120					125			
CTA	ATT	GTC	CAT	TCT	GCA	CTT	CTG	GTT	CCA	TAT	TTT	TCA	TGG	AAA	TAT	612
Leu	Ile	Val	His	Ser	Ala	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	
		130					135					140				
AGC	CAT	CGC	CGC	CAC	CAT	TCT	AAC	ATA	GGA	TCT	CTC	GAG	CGA	GAC	GAA	660
Ser	His	Arg	Arg	His	His	Ser	Asn	Ile	Gly	Ser	Leu	Glu	Arg	Asp	Glu	
	145					150					155					
GTG	TTC	GTC	CCG	AAA	TCA	AAG	TCG	AAA	ATT	TCA	TGG	TAT	TCT	AAG	TAC	708
Val	Phe	Val	Pro	Lys	Ser	Lys	Ser	Lys	Ile	Ser	Trp	Tyr	Ser	Lys	Tyr	
160				165					170					175		
TTA	AAC	AAC	CCG	CCA	GGT	CGA	GTT	TTG	ACA	CTT	GCT	GCC	ACG	CTC	CTC	756
Leu	Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu	
				180					185					190		
CTT	GGC	TGG	CCT	TTA	TAT	TTA	GCT	TTC	AAT	GTC	TCT	GGT	AGA	CCT	TAC	804
Leu	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	
			195					200					205			
GAT	CGC	TTT	GCT	TGC	CAT	TAT	GAT	CCC	TAT	GGC	CCA	ATA	TTT	TCC	GAA	852
Asp	Arg	Phe	Ala	Cys	His	Tyr	Asp	Pro	Tyr	Gly	Pro	Ile	Phe	Ser	Glu	
		210					215					220				
AGA	GAA	AGG	CTT	CAG	ATT	TAC	ATT	GCT	GAC	CTC	GGA	ATC	TTT	GCC	ACA	900
Arg	Glu	Arg	Leu	Gln	Ile	Tyr	Ile	Ala	Asp	Leu	Gly	Ile	Phe	Ala	Thr	
	225					230					235					
ACG	TTT	GTG	CTT	TAT	CAG	GCT	ACA	ATG	GCA	AAA	GGG	TTG	GCT	TGG	GTA	948
Thr	Phe	Val	Leu	Tyr	Gln	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val	
240					245					250				255		
ATG	CGT	ATC	TAT	GGG	GTG	CCA	TTG	CTT	ATT	GTT	AAC	TGT	TTC	CTT	GTT	996
Met	Arg	Ile	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Cys	Phe	Leu	Val	
				260					265					270		
ATG	ATC	ACA	TAC	TTG	CAG	CAC	ACT	CAC	CCA	GCT	ATT	CCA	CGC	TAT	GGC	1044
Met	Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly	
			275					280					285			
TCA	TCG	GAA	TGG	GAT	TGG	CTC	CGG	GGA	GCA	ATG	GTG	ACT	GTC	GAT	AGA	1092
Ser	Ser	Glu	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Met	Val	Thr	Val	Asp	Arg	
		290					295					300				

GAT TAT GGG GTG TTG AAT AAA GTA TTC CAT AAC ATT GCA GAC ACT CAT	1140
Asp Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His	
305 310 315	
GTA GCT CAT CAT CTC TTT GCT ACA GTG CCA CAT TAC CAT GCA ATG GAG	1188
Val Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu	
320 325 330 335	
GCC ACT AAA GCA ATC AAG CCT ATA ATG GGT GAG TAT TAC CGG TAT GAT	1236
Ala Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp	
340 345 350	
GGT ACC CCA TTT TAC AAG GCA TTG TGG AGG GAG GCA AAG GAG TGC TTG	1284
Gly Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu	
355 360 365	
TTC GTC GAG CCA GAT GAA GGA GCT CCT ACA CAA GGC GTT TTC TGG TAC	1332
Phe Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr	
370 375 380	
CGG AAC AAG TAT TAAAAAAGTG TCATGTAGCC TGCCG	1369
Arg Asn Lys Tyr	
385	

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn Ser	
1 5 10 15	
Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr Lys	
20 25 30	
Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His Cys	
35 40 45	
Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Asn Phe Ala Tyr Asn Phe	
50 55 60	
Cys Leu Ser Phe Leu Ser Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr	
65 70 75 80	
Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe	
85 90 95	
Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly	
100 105 110	

His His Ala Phe Ser Glu Tyr Gln Leu Ala Asp Asp Ile Val Gly Leu  
 115 120 125  
 Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser  
 130 135 140  
 His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val  
 145 150 155 160  
 Phe Val Pro Lys Ser Lys Ser Lys Ile Ser Trp Tyr Ser Lys Tyr Leu  
 165 170 175  
 Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Ala Thr Leu Leu Leu  
 180 185 190  
 Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp  
 195 200 205  
 Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Phe Ser Glu Arg  
 210 215 220  
 Glu Arg Leu Gln Ile Tyr Ile Ala Asp Leu Gly Ile Phe Ala Thr Thr  
 225 230 235 240  
 Phe Val Leu Tyr Gln Ala Thr Met Ala Lys Gly Leu Ala Trp Val Met  
 245 250 255  
 Arg Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Cys Phe Leu Val Met  
 260 265 270  
 Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly Ser  
 275 280 285  
 Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp  
 290 295 300  
 Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val  
 305 310 315 320  
 Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala  
 325 330 335  
 Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly  
 340 345 350  
 Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe  
 355 360 365  
 Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg  
 370 375 380  
 Asn Lys Tyr  
 385

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..23
  - (D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGGTATGCC AYGANTGYGG NCA

23

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) TRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..22
  - (D) CTHET INFORMATION: /product=  
"synthetic  
oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAARTGRTGG CACRTGNGTR TC

22



## (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2973 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Arabidopsis thaliana*
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: pAGF2-6
- (ix) FEATURE:
  - (A) NAME/KEY: exon
  - (B) LOCATION: 433..520
- (ix) FEATURE:
  - (A) NAME/KEY: intron
  - (B) LOCATION: 521..1654
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

ATTCGGTAAT TCCTACATAT TTTAGAGATT AGTTTGAGTT TCCATCCATA CTTTACTAGT    60
GATTATAAAT TTAAAATACG TACTTTTTCGA CTATAAAGTG AAAC TAAGTA AATTAGAACG    120
TGATATTAAA AAGTTAATGT TCACTGTTAT ATTTTTTTTCA CAAGTAAAAA ATGGGTTATT    180
TGCGGTAAAT AAAAATACCA GATATTTTGA ATTGATTAAA AAGGTTGAAA TAAGAGAGGA    240
GGGGAAAGAA AAGAAGGTGG GGGCCAGTA TGAAAGGGAA AGGTGTCATC AAATCATCTC    300
TCTCTCTCTC TACCTTCGAC CCACGGGCCG TGTCCATTTA AAGCCCTGTC TCTTGCCATT    360
CCCCATCTGA CCACCAGAAG AAGAGCCACA CACTCACAAA TTAAAAAGAG AGAGAGAGAG    420
AGAGAGACAG AGAGAGAGAG AGATTCTGCG GAGGAGCTTC TTCTTCGTAG GGTGTTTCATC    480
GTTATTAACG TTATCGCCCC TACGTCAGCT CCATCTCCAG GTCCGTCGCT TCTCTTCCAT    540
TTCTTCTCAT TTTGATTTT GATTCTTATT TCTTTCCAGT AGCTCCTGCT CTGTGAATTT    600
CTCCGCTCAC GATAGATCTG CTTATACTCC TTACATTCAA CCTTAGATCT GGTCTCGATT    660
CTCTGTTTCT CTGTTTTTTT CTTTGGTCG AGAATCTGAT GTTTGTTTAT GTTCTGTCAC    720
CATTAATAAT GATGAACTCT CTCATTCATA CAATGATTAG TTTCTCTCGT CTACCAAACG    780

```

ATATGTTGCA	TTTTCACTTT	TCTTCTTTTT	TTCTAAGATG	ATTTGCTTTG	ACCAATTTGT	840
TTAGATCTTT	ATTTTATTTT	ATTTTCTGGT	GGGTTGGTGG	AAATTGAAAA	AAAAAAAAAA	900
AAAAGCATAA	ATTGTTATTT	GTAAATGTAT	TCATTTTTTG	GCTATTTGTT	CTGGGTAAAA	960
ATCTGCTTCT	ACTGTTGAAT	CTTTCCTGGA	TTTTTTACTC	CTATTGGGTT	TTTATAGTAA	1020
AAATACATAA	TAAAAGGAAA	ACAAAAGTTT	TATAGATTCT	CTTAAACCCC	TTACGATAAA	1080
AGTTGGAATC	AAAATAATTC	AGGATCAGAT	GCTCTTTGAT	TGATTCAGAT	GCGATTACAG	1140
TTGCATGGAA	AATTTTCTAG	ATCCGTCGTC	ACATTTTATT	TTCTGTTTAA	ATATCTAAAT	1200
CTGATATATG	ATGTCGACAA	ATTCTGGTGG	CTTATACATC	ACTTCAACTG	TTTTCTTTTG	1260
GCTTTGTTTG	TCAACTTGGT	TTTCAATACG	ATTTGTGATT	TCGATCGCTG	AATTTTTAAT	1320
ACAAGCAAAC	TGATGTTAAC	CACAAGCAAG	AGATGTGACC	TGCCTTATTA	ACATCGTATT	1380
ACTTACTACT	AGTCGTATTC	TCAACGCAAT	CGTTTTTGTA	TTTCTCACAT	TATGCCGCTT	1440
CTCTACTCTT	TATTCCTTTT	GGTCCACGCA	TTTTCTATTT	GTGGCAATCC	CTTTCACAAC	1500
CTGATTTCCC	ACTTTGGATC	ATTTGTCTGA	AGACTCTCTT	GAATCGTTAC	CACTTGTTTC	1560
TTGTGCATGC	TCTGTTTTTT	AGAATTAATG	ATAAACTAT	TCCATAGTCT	TGAGTTTTCA	1620
GCTTGTTGAT	TCTTTTGCTT	TTGGTTTTCT	GCAGAAACAT	GGGTGCAGGT	GGAAGAATGC	1680
CGGTTCCCTAC	TTCTTCCAAG	AAATCGGAAA	CCGACACCAC	AAAGCGTGTG	CCGTGCGAGA	1740
AACCGCCTTT	CTCGGTGGGA	GATCTGAAGA	AAGCAATCCC	GCCGCATTGT	TTCAAACGCT	1800
CAATCCCTCG	CTCTTTCTCC	TACCTTATCA	GTGACATCAT	TATAGCCTCA	TGCTTCTACT	1860
ACGTCGCCAC	CAATTACTTC	TCTCTCCTCC	CTCAGCCTCT	CTCTTACTTG	GCTTGGCCAC	1920
TCTATTGGGC	CTGTCAAGGC	TGTGTCCTAA	CTGGTATCTG	GGTCATAGCC	CACGAATGCG	1980
GTCACCACGC	ATTCAGCGAC	TACCAATGGC	TGGATGACAC	AGTTGGTCTT	ATCTTCCATT	2040
CCTTCCTCCT	CGTCCCTTAC	TTCTCCTGGA	AGTATAGTCA	TCGCCGTCAC	CATTCCAACA	2100
CTGGATCCCT	CGAAAGAGAT	GAAGTATTTG	TCCCAAAGCA	GAAATCAGCA	ATCAAGTGGT	2160
ACGGGAAATA	GCTCAACAAC	GCTCTTGGAC	GCATCATGAT	GTTAACCCTC	CAGTTTCTCC	2220
TCGGGTGGCC	CTTGTACTTA	GCCTTTAACG	TCTCTGGCAG	ACCGTATGAC	GGGTTGGCTT	2280
GCCATTTCTT	CCCCAACGCT	CCCATCTACA	ATGACCGAGA	ACGCCTCCAG	ATATACCTCT	2340
CTGATGCGGG	TATTCTAGCC	GTCTGTTTTG	GTCTTTACCG	TTACGCTGCT	GCACAAGGGA	2400
TGGCCTCGAT	GATCTGCCTC	TACGGAGTAC	CGCTTCTGAT	AGTGAATGCG	TTCCTCGTCT	2460
TGATCACTTA	CTTGCAGCAC	ACTCATCCCT	CGTTGCCTCA	CTACGATTCA	TCAGAGTGGG	2520

ACTGGCTCAG GGGAGCTTTG GCTACCGTAG ACAGAGACTA CGGAATCTTG AACAAAGGTGT 2580  
 TCCACAACAT TACAGACACA CACGTGGCTC ATCACCTGTT CTCGACAATG CCGCATTATA 2640  
 ACGCAATGGA AGCTACAAAG GCGATAAAGC CAATTCTGGG AGACTATTAC CAGTTCGATG 2700  
 GAACACCGTG GTATGTGGCG ATGTATAGGG AGGCAAAGGA GTGTATCTAT GTAGAACCGG 2760  
 ACAGGGAAGG TGACAAGAAA GGTGTGTACT GGTACAACAA TAAGTTATGA GGATGATGGT 2820  
 GAAGAAATTG TCGACTTTTC TCTGTCTGT TTGTCTTTTG TTAAAGAAGC TATGCTTCGT 2880  
 TTTAATAATC TTATTGTCCA TTTTGTGTG TTATGACATT TTGGCTGCTC ATTATGTTAT 2940  
 GTGGGAAGTT AGCGTTCAAA TGT TTTGGGT CGG 2973

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGCATGTNG ARAANARRTG RTG 23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..23

(D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGGCATGTRC TRAAANARRTG RTG

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: JONATHAN EDWARD LIGHTNER  
JOHN JOSEPH OKULEY

(ii) TITLE OF INVENTION: GENES FOR MICROSOMAL  
FATTY ACID DELTA-12  
DESATURASES AND RELATED  
ENZYMES FROM PLANTS

(iii) NUMBER OF SEQUENCES: 17

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: E. I. DU PONT DE NEMOURS  
AND COMPANY  
(B) STREET: 1007 MARKET STREET  
(C) CITY: WILMINGTON  
(D) STATE: DELAWARE  
(E) COUNTRY: U.S.A.  
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: MacIntosh  
(C) OPERATING SYSTEM: Microsoft Windows 95  
(D) SOFTWARE: Microsoft Word Version 7.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: U.S. 07/977,339  
(B) FILING DATE: 17-NOV-1992

(A) APPLICATION NUMBER: U.S. 09/133,962  
(B) FILING DATE: 14-AUG-1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Christenbury, Lynne M..  
(B) REGISTRATION NUMBER: 30,971  
(C) REFERENCE/DOCKET NUMBER: BB1043 US DIV

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (302) 992-5481  
(B) TELEFAX: (302) 892-1026  
(C) TELEX: 235430

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1372 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: p92103

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 93..1244

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGAGAGAGAG ATTCTGCGGA GGAGCTTCTT CTTCGTAGGG TGTTTCATCGT TATTAACGTT  
 60

ATCGCCCCCTA CGTCAGCTCC ATCTCCAGAA AC ATG GGT GCA GGT GGA AGA ATG  
 113

Met Gly Ala Gly Gly Arg Met  
 1 5

CCG GTT CCT ACT TCT TCC AAG AAA TCG GAA ACC GAC ACC ACA AAG CGT  
 161

Pro Val Pro Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr Thr Lys Arg  
 10 15 20

GTG CCG TGC GAG AAA CCG CCT TTC TCG GTG GGA GAT CTG AAG AAA GCA  
 209

Val Pro Cys Glu Lys Pro Pro Phe Ser Val Gly Asp Leu Lys Lys Ala  
 25 30 35

ATC CCG CCG CAT TGT TTC AAA CCG TCA ATC CCT CGC TCT TTC TCC TAC  
 257

Ile Pro Pro His Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ser Tyr  
 40 45 50 55

CTT ATT AGT GAC AAT ATT ATA GGC TCA TGC TTC TAC TAC GTC GGC ACC  
 305

Leu Ile Ser Asp Ile Ile Ile Ala Ser Cys Phe Tyr Tyr Ser Ala Thr  
 60 65 70

AAT TAC TTC TCT CTC CTC CCT CAG CCT CTC TCT TAC TTG GCT TGG CCA  
 353

Asn Tyr Phe Ser Leu Leu Pro Gln Pro Leu Ser Tyr Leu Ala Trp Pro  
 75 80 85

CTC TAT TGG GGC TGT CAA GGC TGT GTC CTA ACT GGT ATC TGG GTC ATA  
 401

```

Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Thr Gly Ile Trp Val Ile
    90                      95                100

GCC CAC GAA TGC GGT CAC CAC GCA TTC AGC GAC TAC CAA TGG CTG GAT
449
Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu Asp
    105                      110                115

GAC ACA GTT GGT CTT ATC TTC CAT TCC TTC CTC CTC GTC CCT TAC TTC
497
Asp Thr Val Gly Leu Ile Phe His Ser Phe Leu Leu Val Pro Tyr Phe
    120                      125                130

TCC TGG AAG TAT AGT CAT CGC CGT CAC CAT TCC AAC ACT GGA TCC CTC
545
Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu
    140                      145                150

GAA AGA GAT GAA GTA TTT GTC CCA AAG CAG AAA TCA GCA ATC AAG TGG
593
Glu Arg Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Ile Lys Trp
    155                      160                165

TAC GGG AAA TAC CTC AAC AAC CCT CTT GGA CGC ATC ATG ATG TTA ACC
641
Tyr Gly Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met Met Leu Thr
    170                      175                180

GTC CAG TTT GTC CTC GGG TGG CCC TTG TAC TTA GGC TTT AAC GTC TCT
689
Val Gln Phe Val Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser
    185                      190                195

GGC AGA CCG TAT GAC GGG TTC GGT TGC CAT TTC TTC CCC AAC GCT CCC
737
Gly Arg Pro Tyr Asp Gly Phe Ala Cys His Phe Phe Pro Asn Ala Pro
    200                      205                210

ATC TAC AAT GAC CGA GAA CGC CTC CAG ATA TAC CTC TCT GAT GCG GGT
785
Ile Tyr Asn Asp Arg Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala Gly
    220                      225                230

ATT CTA GGC GTC TGT TTT GST CTT TAC CGT TAC GGT GGT GCA CAA GGG
833
Ile Leu Ala Val Cys Phe Gly Leu Tyr Arg Tyr Ala Ala Ala Gln Gly
    240                      245                250

ATG GCG TCG ATG ATG TGC CTC TAC GGA GTA CCG CTT CTG ATA GCG AAT
881
Met Ala Ser Met Ile Cys Leu Tyr Gly Val Pro Leu Leu Ile Val Asn
    255                      260                265

GGG TTC CTC GTC TTG ATC ACT TAC TTG CAG CAC ACT CAT CCC TCG TTG
929
Ala Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu

```

```

265          270          275
OCT CAC TAC GAT TCA TCA GAG TGG GAC TGG CTC AGG GGA GCT TTG GCT
977
Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala
280          285          290          295

ACC GTA GAC AGA GAC TAC GGA ATC TTG AAC AAG GTG TTC CAC AAC ATT
1025
Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile
300          305          310

ACA GAC ACA CAC GTG GCT CAT CAC CTG TTC TCG ACA ATG CCG CAT TAT
1073
Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr
315          320          325

AAC GCA ATG GAA GCT ACA AAG GCG ATA AAG CCA ATT CTG GGA GAC TAT
1121
Asn Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Asp Tyr
330          335          340

TAC CAG TTC GAT GGA ACA CCG TGG TAT GTA GCG ATG TAT AGG GAG GCA
1169
Tyr Gln Phe Asp Gly Thr Pro Trp Tyr Val Ala Met Tyr Arg Glu Ala
345          350          355

AAG GAG TGT ATC TAT GTA GAA CCG GAC AGG GAA GGT GAC AAG AAA GGT
1217
Lys Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu Gly Asp Lys Lys Gly
360          365          370          375

GTG TAC TGG TAC AAC AAT AAG TTA TGAGCATGAT GGTGAAGAAA TTGTCGACCT
1271
Val Tyr Trp Tyr Asn Asn Lys Leu
380

TTCTCTTGTC TGTTTGTCTT TTGTTAAAGA AGCTATGCTT CGTTTAAATA ATCTTATTGT
1331

CCATTTTGTT GTGTTATGAC ATTTTGGGTG CTCATTATGT T
1372

```

## (i) INFORMATION FOR SEQ ID NO:2:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38: amino acids  
 (B) TYPE: amino acid  
 (C) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
1          5          10          15

```



Glu	Thr	Asp	Thr	Thr	Lys	Arg	Val	Pro	Cys	Glu	Lys	Pro	Pro	Phe	Ser	
			20					25					30			
Val	Gly	Asp	Leu	Lys	Lys	Ala	Ile	Pro	Pro	His	Cys	Phe	Lys	Arg	Ser	
		35					40					45				
Ile	Pro	Arg	Ser	Phe	Ser	Tyr	Leu	Ile	Ser	Asp	Ile	Ile	Ile	Ala	Ser	
	50					55					60					
Cys	Phe	Tyr	Tyr	Val	Ala	Thr	Asn	Tyr	Phe	Ser	Leu	Leu	Pro	Gln	Pro	
65					70					75					80	
Leu	Ser	Tyr	Leu	Ala	Trp	Pro	Leu	Tyr	Trp	Ala	Cys	Gln	Gly	Cys	Val	
				85					90					95		
Leu	Thr	Gly	Ile	Trp	Val	Ile	Ala	His	Glu	Cys	Gly	His	His	Ala	Phe	
			100					105					110			
Ser	Asp	Tyr	Gln	Trp	Leu	Asp	Asp	Thr	Val	Gly	Leu	Ile	Phe	His	Ser	
	115					120						125				
Phe	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	His	Arg	Arg	His	
130						135					140					
His	Ser	Asn	Thr	Gly	Ser	Leu	Glu	Arg	Asp	Glu	Val	Phe	Val	Pro	Lys	
145					150				155						160	
Gln	Lys	Ser	Ala	Ile	Lys	Trp	Tyr	Gly	Lys	Tyr	Leu	Asn	Asn	Pro	Leu	
				165					170					175		
Gly	Arg	Ile	Met	Met	Leu	Thr	Val	Gln	Phe	Val	Leu	Gly	Trp	Pro	Leu	
			180					185					190			
Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly	Phe	Ala	Cys	
		195					200					205				
His	Phe	Phe	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg	Glu	Arg	Leu	Gln	
	210					215					220					
Ile	Tyr	Leu	Ser	Asp	Ala	Gly	Ile	Leu	Ala	Val	Cys	Phe	Gly	Leu	Tyr	
225					230					235					240	
Arg	Tyr	Ala	Ala	Ala	Gln	Gly	Met	Ala	Ser	Met	Ile	Cys	Leu	Tyr	Gly	
				245					250					255		
Val	Pro	Leu	Leu	Ile	Val	Asn	Ala	Phe	Leu	Val	Leu	Ile	Thr	Tyr	Leu	
			260					265					270			
Gln	His	Thr	His	Pro	Ser	Leu	Pro	His	Tyr	Asp	Ser	Ser	Glu	Trp	Asp	
		275					280					285				
Trp	Leu	Arg	Gly	Ala	Leu	Ala	Thr	Val	Asp	Arg	Asp	Tyr	Gly	Ile	Leu	
	290					295					300					
Asn	Lys	Val	Phe	His	Asn	Ile	Thr	Asp	Thr	His	Val	Ala	His	His	Leu	

305		310		315		320
Phe Ser Thr Met	Pro His Tyr Asn Ala	Met Glu Ala Thr	Lys Ala Ile			
	325	330	335			
Lys Pro Ile Leu	Gly Asp Tyr Tyr	Gln Phe Asp Gly Thr	Pro Trp Tyr			
	340	345	350			
Val Ala Met Tyr Arg	Glu Ala Lys Glu Cys Ile	Tyr Val Glu Pro Asp				
	355	360	365			
Arg Glu Gly Asp Lys Lys	Gly Val Tyr Trp Tyr	Asn Asn Lys Leu				
	370	375	380			

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1426 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA to mRNA

## (iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 130..1284

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GGCACGAGCT CGTGCCGAAT TCGGCACGAG AGGAGACAGA GAGAGAGTTT GAGGAGGAGC   60
TTCTTCGTAG GGTTCATCGT TATTAACGTT AAATCTTCAT CCCCCCCTAC GTCAGCCAGC  120
TCAAGAAAC ATG GGT GCA GGT GGA AGA ATG CAA GTG TCT CCT CCC TCC   168
    Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser
        1             5             10
AAA AAG TCT GAA ACC GAC AAC ATC AAG CGC GTA CCC TGC GAG ACA CGG   216
Lys Lys Ser Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro
    15             20             25
CCC TTC ACT GTC GGA GAA CFC AAG AAA GGA ATC CCA CCG CAC TGT TTC   264
Pro Phe Thr Val Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe
    30             35             40
AAA CGC TCG ATC CCT CGC TCT TTC TCC TAC CTC ATC TGG GAC ATC ATC   312
Lys Arg Ser Ile Pro Arg Ser Phe Ser Tyr Leu Ile Trp Asp Ile Ile
        50             55             60
ATA GGC TGC TGC TTC TAC TAC GTC GGC ACC ACT TAC TTC CCT CTC CTC   360

```

Ile	Ala	Ser	Cys	Phe	Tyr	Tyr	Val	Ala	Thr	Thr	Tyr	Phe	Pro	Leu	Leu	
			65					70					75			
CCT	CAC	CCT	CTC	TCC	TAC	TTC	GCC	TGG	CCT	CTC	TAC	TGG	GCC	TGC	CAG	408
Pro	His	Pro	Leu	Ser	Tyr	Phe	Ala	Trp	Pro	Leu	Tyr	Trp	Ala	Cys	Gln	
		80					85					90				
GGC	TGC	GTC	CTA	ACC	GGC	GTC	TGG	GTC	ATA	GCC	CAC	GAG	TGC	GGC	CAC	456
Gly	Cys	Val	Leu	Thr	Gly	Val	Trp	Val	Ile	Ala	His	Glu	Cys	Gly	His	
	95					100					105					
CAC	GCC	TTC	AGC	GAC	TAC	CAG	TGG	CTG	GAC	GAC	ACC	GTC	GGC	CTC	ATC	504
His	Ala	Phe	Ser	Asp	Tyr	Gln	Trp	Leu	Asp	Asp	Thr	Val	Gly	Leu	Ile	
110					115					120					125	
TTC	CAC	TCC	TTC	CTC	CTC	GTC	CCT	TAC	TTC	TCC	TGG	AAG	TAC	AGT	CAT	552
Phe	His	Ser	Phe	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	His	
				130					135					140		
CGA	CGC	CAC	CAT	TCC	AAC	ACT	GGC	TCC	CTC	GAG	AGA	GAC	GAA	GTG	TTT	600
Arg	Arg	His	His	Ser	Asn	Thr	Gly	Ser	Leu	Glu	Arg	Asp	Glu	Val	Phe	
				145				150					155			
GTC	CCC	AAG	AAG	AAG	TCA	GAC	ATC	AAG	TGG	TAC	GGC	AAG	TAC	CTC	AAC	648
Val	Pro	Lys	Lys	Lys	Ser	Asp	Ile	Lys	Trp	Tyr	Gly	Lys	Tyr	Leu	Asn	
		160					165					170				
AAC	CCG	TTG	GGA	CGC	ACC	GTG	ATG	TTA	ACG	GTT	CAG	TTC	ACT	CTC	GGC	696
Asn	Pro	Leu	Gly	Arg	Thr	Val	Met	Leu	Thr	Val	Gln	Phe	Thr	Leu	Gly	
	175					180					185					
TGG	CCT	TTG	TAC	TTA	GCC	TTC	AAC	GTC	TGC	GGG	AGA	CCT	TAC	GAC	GGC	744
Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly	
190					195					200					205	
GGC	TTG	GCT	TGC	CAT	TTC	CAC	CCC	AAC	GCT	CCC	ATC	TAC	AAC	GAC	CGT	792
Gly	Phe	Ala	Cys	His	Phe	His	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg	
				210					215					220		
GAG	CGT	CTC	CAG	ATA	TAC	ATC	TCC	GAC	GCT	GGC	ATC	CTC	GCC	GTC	TGC	840
Glu	Arg	Leu	Gln	Ile	Tyr	Ile	Ser	Asp	Ala	Gly	Ile	Leu	Ala	Val	Cys	
			225					230					235			
TAC	GGT	CTC	TAC	CGC	TAC	GCT	GCT	GTC	CAA	GGA	GTT	GCC	TGC	ATG	GTC	888
Tyr	Gly	Leu	Tyr	Arg	Tyr	Ala	Ala	Val	Gln	Gly	Val	Ala	Ser	Met	Val	
		24				245						250				
TGC	TTC	TAC	GGA	GTT	CCT	CTT	CTG	ATT	GTC	AAC	GGG	TTC	TTA	GTT	TTG	936
	Phe	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Gly	Phe	Leu	Val	Leu	
	255					260					265					
ATC	ACT	TAC	TTG	CAG	CAC	ACG	CAT	CCT	TCC	CTG	CCT	CAC	TAT	GAC	TGC	984
Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ser	Leu	Pro	His	Tyr	Asp	Ser	
270					275					280					285	
TCT	GAG	TGG	GAT	TGG	TTG	AGG	GGA	GCT	TTG	GCC	ACC	GTT	GAC	AGA	GAC	1032

```

Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp
      290      295      300
TAC GGA ATC TTG AAC AAG GTC TTC CAC AAT ATC ACG GAC ACG CAC GTG 1080
Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val
      305      310      315
GCG CAT CAC CTG TTC TCG ACC ATG CCG CAT TAT CAT GCG ATG GAA GCT 1128
Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala
      320      325      330
ACG AAG GCG ATA AAG CCG ATA CTG GGA GAG TAT TAT CAG TTC GAT GGG 1176
Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly
      335      340      345
ACG CCG GTG GTT AAG GCG ATG TGG AGG GAG GCG AAG GAG TGT ATC TAT 12
24
Thr Pro Val Val Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr
350      355      360      365
GTG GAA CCG GAC AGG CAA GGT GAG AAG AAA GGT GTG TTC TGG TAC AAC 12
72
Val Glu Pro Asp Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn
      370      375      380
AAT AAG TTA TGA AGCAAAGAAG AAAGTGAACC TTTCTCTTCT ATGATGTGTCT 13
24
Asn Lys Leu
TTGTTTAAGA AGCTATGTTT CTGTTTCAAT AATCTTAATT ATCCATTTTG TTGTGTTTTC 13
84
TGACATTTTG GCTAAAATTA TGTGATGTTG GAAGTTAGTG TC 14
26

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser Lys Lys Ser
  1      5      10
Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro Pro Phe Thr
      20      25      30
Val Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
      35      40      45

```

Ile	Pro	Arg	Ser	Phe	Ser	Tyr	Leu	Ile	Trp	Asp	Ile	Ile	Ile	Ala	Ser
50						55					60				
Cys	Phe	Tyr	Tyr	Val	Ala	Thr	Thr	Tyr	Phe	Pro	Leu	Leu	Pro	His	Pro
65					70					75					80
Leu	Ser	Tyr	Phe	Ala	Trp	Pro	Leu	Tyr	Trp	Ala	Cys	Gln	Gly	Cys	Val
				85					90					95	
Leu	Thr	Gly	Val	Trp	Val	Ile	Ala	His	Glu	Cys	Gly	His	His	Ala	Phe
			100					105					110		
Ser	Asp	Tyr	Gln	Trp	Leu	Asp	Asp	Thr	Val	Gly	Leu	Ile	Phe	His	Ser
		115					120					125			
Phe	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	His	Arg	Arg	His
	130					135					140				
His	Ser	Asn	Thr	Gly	Ser	Leu	Glu	Arg	Asp	Glu	Val	Phe	Val	Pro	Lys
145					150					155					160
Lys	Lys	Ser	Asp	Ile	Lys	Trp	Tyr	Gly	Lys	Tyr	Leu	Asn	Asn	Pro	Leu
				165					170					175	
Gly	Arg	Thr	Val	Met	Leu	Thr	Val	Gln	Phe	Thr	Leu	Gly	Trp	Pro	Leu
			180					185					190		
Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly	Gly	Phe	Ala
		195					200					205			
Cys	His	Phe	His	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg	Glu	Arg	Leu
	210					215					220				
Gln	Ile	Tyr	Ile	Ser	Asp	Ala	Gly	Ile	Leu	Ala	Val	Cys	Tyr	Gly	Leu
225					230					235					240
Tyr	Arg	Tyr	Ala	Ala	Val	Gln	Gly	Val	Ala	Ser	Met	Val	Cys	Phe	Tyr
				245					250					255	
Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Gly	Phe	Leu	Val	Leu	Ile	Thr	Tyr
			260					265					270		
Leu	Gln	His	Thr	His	Pro	Ser	Leu	Pro	His	Tyr	Asp	Ser	Ser	Glu	Trp
		275					280					285			
Asp	Trp	Leu	Arg	Gly	Ala	Leu	Ala	Thr	Val	Asp	Arg	Asp	Tyr	Gly	Ile
	290					295					300				
Leu	Asn	Lys	Val	Phe	His	Asn	Ile	Thr	Asp	Thr	His	Val	Ala	His	His
305					310					315					320
Leu	Phe	Ser	Thr	Met	Pro	His	Tyr	His	Ala	Met	Glu	Ala	Thr	Lys	Ala
				325					330					335	
Ile	Lys	Pro	Ile	Leu	Gly	Glu	Tyr	Tyr	Gln	Phe	Asp	Gly	Thr	Pro	Val
			340					345					350		

Val Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro  
 355 360 365  
 Asp Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu  
 370 375 380

## (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1462 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: pSF2-165F

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 108..1247

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATATACTA ATATTTGCTT GTATTGATAG CCCCTCCGTT CCCAAGAGTA TAAAACTGCA  
 60

TCGAATAATA CAAGCCACTA GGCATGGGTC TAGCAAAGGA AACAACA ATG GGA GGT  
 116

Met Gly Gly  
 1

AGA GGT CGT GTG GGC AAA GTG GAA GTT CAA GGG AAG AAG CCT CTC TCA  
 164

Arg Gly Arg Val Ala Lys Val Gln Val Gln Gly Lys Lys Pro Leu Ser  
 5 10 15

ATG GAT CCA AAG ACA AAG CCA CCA TTC ACT GTT GGC CAA CTC AAG AAA  
 212

Arg Val Pro Asn Thr Lys Phe Phe Phe Thr Val Gly Gln Leu Lys Lys  
 20 25 30 35

GCA ATT CCA CCA CAA TAC TTT CAG CCG TCC CTC CTC ACT TCA TTC TCC  
 260

Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr Ser Phe Ser  
 40 45 50

TAT GTT GTT TAT GAC CTT TCA TTT GCC TTC ATT TTC TAC ATT GCC ACC  
 308  
 Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr Ile Ala Thr  
 55 60 65

ACC TAC TTC CAC CTC CTT CCT CAA CCC TTT TCC CTC ATT GCA TGG CCA  
 356  
 Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile Ala Trp Pro  
 70 75 80

ATC TAT TGG GTT CTC CAA GGT TGC CTT CTC ACT GGT GTG TGG GTG ATT  
 404  
 Ile Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr Gly Val Trp Val Ile  
 85 90 95

GCT CAC GAG TGT GGT CAC CAT GCC TTC AGC AAG TAC CAA TGG GTT GAT  
 452  
 Ala His Glu Cys Gly His His Ala Phe Ser Lys Tyr Gln Trp Val Asp  
 100 105 110 115

GAT GTT GTG GGT TTG ACC CTT CAC TCA ACA CTT TTA GTC CCT TAT TTC  
 500  
 Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val Pro Tyr Phe  
 120 125 130

TCA TGG AAA ATA AGC CAT CGC CGC CAT CAC TCC AAC ACA GGT TCC CTT  
 548  
 Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu  
 135 140 145

GAC CGT GAT GAA GTG TTT GTC CCA AAA CCA AAA TCC AAA GTT GCA TGG  
 596  
 Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp  
 150 155 160

TTT TCC AAG TAC TTA AAC AAC CCT CTA GGA AGG GCT GTT TCT CTT CTC  
 644  
 Phe Ser Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val Ser Leu Leu  
 165 170 175

GTC ACA CTC ACA ATA GGG TGG CCT ATG TAT TTA GCC TTC AAT GTC TCT  
 692  
 Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser  
 180 185 190 195

GCG AGA CCC CAT GAT AGT TTT GCA AAG CAC TAC CAC CCT TAT GGT CCC  
 740  
 Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro Tyr Ala Pro  
 200 205 210

ATA TAT TCT AAC CGT GAG AGG CTT CTG ATC TAT GTC TCT CAT GTT GCT  
 788  
 Ile Tyr Ser Asn Arg Glu Arg Leu Leu Ile Tyr Val Ser Asp Val Ala  
 215 220 225

TTG TTT TCT GTG AAT TAC TCT CTC TAC CTT GTT GCA ACC CTG AAA GGG

836  
 Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr Leu Lys Gly  
           230                          235                          240

TTG GTT TGG CTG CTA TGT GTT TAT GGG GTG CCT TTG CTC ATT GTG AAC  
 884  
 Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn  
           245                          250                          255

GGT TTT CTT GTG ACT ATC ACA TAT TTG CAG CAC ACA CAC TTT GCC TTG  
 932  
 Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu  
           260                          265                          270                          275

CCT CAT TAC GAT TCA TCA GAA TGG GAC TGG CTG AAG GGA GCT TTG GCA  
 980  
 Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala  
                                   280                          285                          290

ACT ATG GAC AGA GAT TAT GGG ATT CTG AAC AAG GTG TTT CAT CAC ATA  
 1028  
 Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His His Ile  
                           295                          300                          305

ACT GAT ACT CAT GTG GCT CAC CAT CTC TTC TCT ACA ATG CCA CAT TAC  
 1076  
 Phe Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr  
           310                          315                          320

CAT GCA ATG GAG GCA ACC AAT GCA ATC AAG CCA ATA TTG GGT GAG TAC  
 1124  
 His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr  
           325                          330                          335

TAC CAA TTT GAT GAC ACA CCA TTT TAC AAG GCA CTG TGG AGA GAA GCG  
 1172  
 Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala  
           340                          345                          350                          355

AGA GAG TGC CTC TAT GTG GAG CCA GAT GAA GGA ACA TCC GAG AAG GGC  
 1220  
 Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly  
                           360                          365                          370

GTG TAT TGG TAC AGG AAC AAG TAT TGATGGAGCA ACCAATGGGC CATAGTGGGA  
 1274  
 Val Tyr Trp Tyr Arg Asn Lys Tyr                          380

GTTATGGAAG TTTTGTGATG TATTAGTACA TAATTAGTAG AATGTTATAA ATAAGTGGAT  
 1324

TTGCGCGTA ATGACTTTGT GTSTATTGTG AAACAGCTTG TTGCGATCAT GGTATAATG  
 1374

TAAAAATAAT TCTGGTATTA ATTACATGTG GAAAGTCTT TATTATAGG TTTCTGCGTA



1454

AAAAAAA

1462

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 379 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Gly Gly Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys
 1           5           10           15

Pro Leu Ser Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln
 20           25           30

Leu Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr
 35           40           45

Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr
 50           55           60

Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile
 65           70           75           80

Ala Trp Pro Ile Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr Gly Val
 85           90           95

Trp Val Ile Ala His Glu Cys Gly His His Ala Phe Ser Lys Tyr Gln
100           105           110

Trp Val Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val
115           120           125

Pro Tyr Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr
130           135           140

Gly Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys
145           150           155           160

Val Ala Trp Thr Ser Lys Tyr Leu His Asn Thr Leu Gly Arg Ala Val
165           170           175

Ser Leu Leu Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe
180           185           190

Asn Val Ser Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro
195           200           205

Tyr Ala Pro Ile Tyr Ser Asn Arg Glu Asn Leu Leu Ile Tyr Val Ser

```

210	215	220
Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr 225 230 235 240		
Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu 245 250 255		
Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His 260 265 270		
Phe Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly 275 280 285		
Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe 290 295 300		
His His Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met 305 310 315 320		
Pro His Tyr His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu 325 330 335		
Gly Glu Tyr Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp 340 345 350		
Arg Glu Ala Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser 355 360 365		
Glu Lys Gly Val Tyr Trp Tyr Arg Asn Lys Tyr 370 375		

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1730 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (v) ORGANISM: Zea mays

## (vi) ORGANISM: Zea mays

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: pFad.#1

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 165..1328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGCGCTCTCC CCTCCCTCCT CCCTGCAAAT CCTGCAGACA CCACCGGTCCG TTTTCTCTC  
60

CGGGACAGGA GAAAAGGGGA GAGAGAGGTG AGGCGGGGTG TCCGCCCGAT CTGCTCTGCC  
120

CGGACGCAGC TGTTACGACC TCCTCAGTCT CAGTCAGGAG CAAG ATG GGT GCC GGC  
175

Met Gly Ala Gly  
1

GGC AGG ATG ACC GAG AAG GAG CGG GAG AAG CAG GAG CAG CTC GCC CGA  
224

Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu Gln Leu Ala Arg  
5 10 15 20

GCT ACC GGT GGC GCC GCG ATG CAG CGG TCG CCG GTG GAG AAG CCT CCG  
272

Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val Glu Lys Pro Pro  
25 30 35

TTC ACT CTG GGT CAG ATC AAG AAG GGC ATC CCG CCA CAC TGC TTC GAG  
320

Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro His Cys Phe Glu  
40 45 50

CGC TCG GTG CTC AAG TCC TTC TCG TAC GTG GTC CAC GAC CTG GTG ATC  
368

Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile  
55 60 65

GCC GCG GCG CTC CTC TAC TTC GCG CTG GCC ATC ATA CCG GCG CTC CCA  
416

Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile Pro Ala Leu Pro  
70 75 80

AGC CCG CTC CGC TAC GCC GCC TGG CCG CTG TAC TGG ATC GCG CAG GGG  
464

Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp Ile Ala Gln Gly  
85 90 95 100

TTC GTG TGC ACC GGC GTG TGG GTC ATC GCG CAC GAG TGC GGC CAC CAC  
512

Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His  
105 110 115

GCC TTC TCG GAC TAC TCG CTC CTG GAC GAC GTG GTC GGC CTG GTG CTC  
560

Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Val Val Gly Leu Val Leu  
120 125 130

CAC TCG TCG CTC ATG GTG CCG TAC TTC TCG TGG AAG TAC AGC CAC CCG  
608

His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg

135	140	145
CGC CAC CAC TCC AAC ACG GGG TCC CTG GAG CGC GAC GAG GTG TTC GTG 656		
Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val 150 155 160		
CCC AAG AAG AAG GAG GCG CTG CCG TGG TAC ACC CCG TAC GTG TAC AAC 704		
Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro Tyr Val Tyr Asn 165 170 175 180		
AAC CCG GTC GGC CGG GTG GTG CAC ATC GTG GTG CAG CTC ACC CTC GGG 752		
Asn Pro Val Gly Arg Val Val His Ile Val Val Gln Leu Thr Leu Gly 185 190 195		
TGG CCG CTG TAC CTG GCG ACC AAC GCG TCG GGG CCG CCG TAC CCG CGC 800		
Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg Pro Tyr Pro Arg 200 205 210		
TTC GCC TGC CAC TTC GAC CCC TAC GGC CCC ATC TAC AAC GAC CGG GAG 848		
Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu 215 220 225		
CGC GCC CAG ATC TTC GTC TCG GAC GCC GGC GTC GTG GCC GTG GCG TTC 896		
Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val Ala Val Ala Phe 230 235 240		
GGG CTG TAC AAG CTG GCG GCG GCG TTC GGG GTC TGG TGG GTG GTG CGC 944		
Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp Trp Val Val Arg 245 250 255 260		
GTG TAC GCC GTG CCG CTG CTG ATC GTG AAC GCG TGG CTG GTG CTC ATC 992		
Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile 265 270 275		
ACC TAC CTG CAG CAC ACC CAC CCG TCG CTC CCC CAC TAC GAC TCG AGC 1040		
Ile Tyr Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser 280 285 290		
GAG TGG GAC TGG CTG CCG GGC GCG CTG GCG ACC ATG GAC CGC GAC TAC 1088		
Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr 295 300 305		
GGC ATC CTC AAC CCG GTG TTC CAC AAC ATC ACG GAC ACG CAC GTC GCG 1136		
Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp Thr His Val Ala 310 315 320		

CAC CAC CTC TTC TCC ACC ATG CCG CAC TAC CAC GCC ATG GAG GCC ACC  
 1184  
 His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr  
 325 330 335 340

AAG GCG ATC AGG CCC ATC CTC GGC GAC TAC TAC CAC TTC GAC CCG ACC  
 1232  
 Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Pro Thr  
 345 350 355

CCT GTC GCC AAG GCG ACC TGG CCG GAG GCC GGG GAA TGC ATC TAC GTC  
 1280  
 Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu Cys Ile Tyr Val  
 360 365 370

GAG CCC GAG GAC CCG AAG GGC GTC TTC TGG TAC AAC AAG AAG TTC TAGCCGCCG  
 C 1335  
 Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn Lys Lys Phe  
 375 380 385

CGCTCGCAGA GCTGAGGACG CTACCGTAGG AATGGGAGCA GAAACCAGGA GGAGGAGACG  
 1395

GTACTCGCCC CAAAGTCTCC GTCAACCTAT CTAATCGTTA GTCGTCAGTC TTTTAGACGG  
 1455

GAAGAGAGAT CATTTGGGCA CAGAGACGAA GCCTTACTGC AGTGCCATCG CTAGAGCTGC  
 1515

CATCAAGTAC AAGTAGGCAA ATTGCTCAAC TTAGTGTGTC CCATGTTGTT TTTCTTAGTC  
 1575

GTCCGCTGCT GTAGGCTTTC CGGCGGGCGT CTTTTGTGTG GTTGGCATCC GTGGCCATGC  
 1635

CTGTGCGTGC GTGGCCGCGC TTGTGCTGTG CGTCTGTGCT GCGGTTGGCG TCGTCTCTTC  
 1695

GTGCTCCCCG TGTGTTGTTG TAAAACAAGA AGATGTTTTT TGGTGTCTTT GGCGGAATAA  
 1755

CAGATCGTCC GAACGAAAAA AAAAAAAAAA AAAAA  
 1790

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met 1	Gly	Ala	Gly	Gly 5	Arg	Met	Thr	Glu	Lys 10	Glu	Arg	Glu	Lys	Gln 15	Glu
Gln	Leu	Ala	Arg 20	Ala	Thr	Gly	Gly	Ala 25	Ala	Met	Gln	Arg	Ser 30	Pro	Val
Glu	Lys	Pro 35	Pro	Phe	Thr	Leu	Gly 40	Gln	Ile	Lys	Lys	Ala 45	Ile	Pro	Pro
His	Cys 50	Phe	Glu	Arg	Ser	Val 55	Leu	Lys	Ser	Phe	Ser 60	Tyr	Val	Val	His
Asp 65	Leu	Val	Ile	Ala	Ala 70	Ala	Leu	Leu	Tyr	Phe 75	Ala	Leu	Ala	Ile	Ile 80
Pro	Ala	Leu	Pro	Ser 85	Pro	Leu	Arg	Tyr	Ala 90	Ala	Trp	Pro	Leu	Tyr 95	Trp
Ile	Ala	Gln	Gly 100	Cys	Val	Cys	Thr	Gly 105	Val	Trp	Val	Ile	Ala 110	His	Glu
Cys	Gly	His 115	His	Ala	Phe	Ser	Asp 120	Tyr	Ser	Leu	Leu	Asp 125	Asp	Val	Val
Gly	Leu	Val 130	Leu	His	Ser	Ser 135	Leu	Met	Val	Pro	Tyr 140	Phe	Ser	Trp	Lys
Tyr 145	Ser	His	Arg	Arg	His 150	His	Ser	Asn	Thr	Gly 155	Ser	Leu	Glu	Arg	Asp 160
Glu	Val	Phe	Val	Pro 165	Lys	Lys	Lys	Glu	Ala 170	Leu	Pro	Trp	Tyr	Thr 175	Pro
Tyr	Val	Tyr	Asn 180	Asn	Pro	Val	Gly	Arg 185	Val	Val	His	Ile	Val 190	Val	Gln
Leu	Thr 195	Leu	Gly	Trp	Pro	Leu	Tyr 200	Leu	Ala	Thr	Asn	Ala 205	Ser	Gly	Arg
Pro	Tyr 210	Pro	Arg	Phe	Ala	Cys 215	His	Phe	Asp	Pro	Tyr 220	Gly	Pro	Ile	Tyr
Asn 225	Asp	Arg	Glu	Arg	Ala 230	Gln	Ile	Phe	Val	Ser 235	Asp	Ala	Gly	Val	Val 240
Ala	Val	Ala	Phe 245	Gly	Leu	Tyr	Leu	Leu	Ala 250	Ala	Ala	Phe	Gly	Val 255	Trp
Trp	Val	Val	Arg 260	Val	Tyr	Ala	Val	Ser 265	Leu	Leu	Ile	Val	Asn 270	Ala	Trp
Leu	Val	Leu 275	Ile	Thr	Tyr	Leu	Gln 280	His	Thr	His	Pro	Ser 285	Leu	Pro	His
Tyr 290	Asp	Ser	Ser	Glu	Trp	Asp 295	Trp	Leu	Arg	Gly	Ala 300	Leu	Ala	Thr	Met

Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp  
 305 310 315 320  
 Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala  
 325 330 335  
 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His  
 340 345 350  
 Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu  
 355 360 365  
 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn  
 370 375 380  
 Lys Lys Phe  
 385

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 673 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Ricinus communis

(vii) IMMEDIATE SOURCE:

(B) CLONE: pRF2-1C

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 1..673

(x) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCC GTC ATC GCG CAG GAT TGT GGG CAC CAT GCC TTC AGT GAC TAT CAA  
 48  
 TAT TTT TAT ATT TTA Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln  
 1 5 10 15  
 TTG CTT GAT GAT GTA GTT GST CTT ATC CTA CAC TCC TGT CTC CTT GTC  
 26  
 Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val  
 20 25 30  
 CCT TAT TTT TCA TGG AAA CAC AGC CAT TGC CGA CAT CAT TCC AAC ACA

144  
 Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr  
           35                          40                          45  
  
 GGG TCC CTG GAA CGG GAT GAA GTG TTT GTT CCC AAG AAG AAA TCT AGT  
 192  
 Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser  
           50                          55                          60  
  
 ATC CGT TGG TAT TCC AAA TAC CTC AAC AAC CCT CCA GGT CGT ATC ATG  
 240  
 Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met  
           65                          70                          75                          80  
  
 ACA ATT GCC GTC ACA CTT TCA CTT GGC TGG CCT CTG TAC CTA GCA TTC  
 288  
 Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe  
                           85                          90                          95  
  
 AAT GTT TCA GGC AGG CCA TAT GAT CGG TTC GGC TGC CAC TAT GAC CCA  
 336  
 Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro  
                           100                          105                          110  
  
 TAT GGC CCG ATC TAC AAT GAT CGC GAG CGA ATC GAG ATA TTC ATA TCA  
 384  
 Thr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser  
                           115                          120                          125  
  
 GAT GCT GGT GTT CTT GCT GTC ACT TTT GGT CTC TAC CAA CTT GCT ATA  
 432  
 Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile  
           130                          135                          140  
  
 GCG AAG GGG CTT GCT TGG GTT GTC TGT GTA TAT GGA GTG CCA TTG TTG  
 480  
 Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu  
           145                          150                          155                          160  
  
 GTG GTG AAT TCA TTC CTT GTT CTG ATC ACA TTT CTG CAG CAT ACT CAC  
 528  
 Val Val Asn Ser Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His  
                           165                          170                          175  
  
 CCG GCA TTG CCA CAT TAT GAT TCG TCG GAG TGG GAC TGG CTA AGA GGA  
 576  
 Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Tyr Asp Phe Leu Arg Gly  
                           180                          185                          190  
  
 GGT CTA GCA ACT GTT GAC AGA GAT TAC GGG ATC TTG AAC AAG GTG TTC  
 624  
 Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe  
                           195                          200                          205  
  
 CAT AAC ATA ACG GAC ACT CAA GTA GCT CAC CAC CTT TTC ACC ATG CCC C  
 672



His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro  
 210 215 220

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 224 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln  
 1 5 10 15  
 Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val  
 20 25 30  
 Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr  
 35 40 45  
 Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser  
 50 55 60  
 Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met  
 65 70 75 80  
 Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe  
 85 90 95  
 Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro  
 100 105 110  
 Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser  
 115 120 125  
 Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile  
 130 135 140  
 Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu  
 145 150 155 160  
 Val Val Asn Ser Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His  
 165 170 175  
 Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Thr Asp Trp Leu Arg Gly  
 180 185 190  
 Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe  
 195 200 205  
 His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro  
 210 215 220

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1369 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Ricinus communis

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: pRF197c-42

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 184..1347

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGGCCGGGAT TCCGGTTTTC ACASTAATTT CCAAAAAATG CATGATTTC CCTCAAATCA  
 60

AACACCACAC CTTATAACTT AGTCTTARGA GAGAGAGAGA GAGGAGACAT TTCTCTTCTC  
 120

TGAGATGAGC ACTTCTCTTC CAGACATCGA AGCCTCAGGA AAGTGCTTGA GAAGAGCTTG  
 180

AGA ATG GGA GGT GGT GGT CGC ATG TCT ACT GTC ATA ATC AGC AAC AAC  
 228

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn  
 1 5 10 15

AGT GAG AAG AAA GGA GGA AGG AGC CAC CTG GAG CGA GCG CCG CAC ACG  
 276

Ser Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr  
 20 25 30

AAG CCG CCG TAT ACA GTT GGT AAT CCG AAG ASA GCG ATC CGA CCG CAT  
 34

Lys Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His  
 35 40 45

TGC TTT GAA CGC TCT TTT GTG CGC TCA TTC TCC AAT TTT GCC TAT AAT  
 372

Cys Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Asn Phe Ala Tyr Asn  
 50 55 60

TTC TGC TTA AGT TTT GTT TGC TAG TCG ATC GCG ACC AAC TTC TTC GGT

420  
 Phe Cys Leu Ser Phe Leu Ser Tyr Ser Ile Ala Thr Asn Phe Phe Pro  
     65                                    70                    75  
  
 TAC ATC TCT TCT CCG CTC TCG TAT GTC GCT TGG CTG GTT TAC TGG CTC  
 468  
 Tyr Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu  
     80                                    85                    90                    95  
  
 TTC CAA GGC TGC ATT CTC ACT GGT CTT TGG GTC ATC GGC CAT GAA TGT  
 516  
 Phe Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys  
                                     100                    105                    110  
  
 GGC CAT CAT GCT TTT AGT GAG TAT CAG CTG GCT GAT GAC ATT GTT GGC  
 564  
 Gly His His Ala Phe Ser Glu Tyr Gln Leu Ala Asp Asp Ile Val Gly  
                                     115                    120                    125  
  
 CTA ATT GTC CAT TCT GCA CTT CTG GTT CCA TAT TTT TCA TGG AAA TAT  
 612  
 Leu Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr  
                                     130                    135                    140  
  
 AGC CAT CGC CGC CAC CAT TCT AAC ATA GGA TCT CTC GAG CGA GAC GAA  
 660  
 Ser His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu  
                                     145                    150                    155  
  
 GTG TTC GTC CCG AAA TCA AAG TCG AAA ATT TCA TGG TAT TCT AAG TAC  
 708  
 Val Phe Val Pro Lys Ser Lys Ser Lys Ile Ser Trp Tyr Ser Lys Tyr  
     160                                    165                    170                    175  
  
 TTA AAC AAC CCG CCA GGT CGA GTT TTG ACA CTT GCT GCC ACG CTC CTC  
 756  
 Leu Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Ala Thr Leu Leu  
                                     180                    185                    190  
  
 CTT GGC TGG CCT TTA TAT TTA GCT TTC AAT GTC TCT GGT AGA CCT TAC  
 804  
 Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr  
                                     195                    200                    205  
  
 GAT CGC TTT GCT TGC CAT TAT GAT CCC TAT GGC CCA ATA TTT TCC GAA  
 852  
 Asp Arg Thr Ala Cys His Tyr Asp Pro Tyr Gly Leu Thr Thr Ser Glu  
                                     210                    215                    220  
  
 AGA GAA AGG CTT CAG ATT TAC ATT GCT GAC CTC GGA ATC TTT GCC ACA  
 900  
 Arg Gln Arg Leu Gln Ile Tyr Ile Ala Asp Leu Gly Ile Phe Ala Thr  
                                     225                    230                    235  
  
 ACG TTT GTG CTT TAT CAG GCT ACA ATG GCA AAA GGG TTG GCT TGG GTA  
 948

```

Thr Phe Val Leu Tyr Gln Ala Thr Met Ala Lys Gly Leu Ala Trp Val
240                               245                250                255

ATG CGT ATC TAT GGG GTG CCA TTG CTT ATT GTT AAC TGT TTC CTT GTT
996
Met Arg Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Cys Phe Leu Val
                260                               265                270

ATG ATC ACA TAC TTG CAG CAC ACT CAC CCA GCT ATT CCA CGC TAT GGC
1044
Met Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly
                275                               280                285

TCA TCG GAA TGG GAT TGG CTC CGG GGA GCA ATG GTG ACT GTC GAT AGA
1092
Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg
                290                               295                300

GAT TAT GGG GTG TTG AAT AAA GTA TTC CAT AAC ATT GCA GAC ACT CAT
1140
Asp Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His
                305                               310                315

GTA GCT CAT CAT CTC TTT GCT ACA GTG CCA CAT TAC CAT GCA ATG GAG
1188
Val Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu
320                               325                330

GCC ACT AAA GCA ATC AAG CCT ATA ATG GGT GAG TAT TAC CGG TAT GAT
1236
Ala Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp
                340                               345                350

GGT ACC CCA TTT TAC AAG GCA TTG TGG AGG GAG GCA AAG GAG TGC TTG
1284
Gly Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu
                355                               360                365

ITC GTC GAG CCA GAT GAA GGA GGT CCT ACA CAA GGC GTT TTC TGG TAC
1332
Phe Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr
                370                               375                380

CGG AAC AAG TAT TAAAAAGTG TCATGTAGCC TGCCG
1369
Arg Asn Lys Tyr

```

2. INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 387 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Gly	Gly	Gly	Gly	Arg	Met	Ser	Thr	Val	Ile	Ile	Ser	Asn	Asn	Ser	1	5	10	15
Glu	Lys	Lys	Gly	Gly	Ser	Ser	His	Leu	Glu	Arg	Ala	Pro	His	Thr	Lys	20	25	30	
Pro	Pro	Tyr	Thr	Leu	Gly	Asn	Leu	Lys	Arg	Ala	Ile	Pro	Pro	His	Cys	35	40	45	
Phe	Glu	Arg	Ser	Phe	Val	Arg	Ser	Phe	Ser	Asn	Phe	Ala	Tyr	Asn	Phe	50	55	60	
Cys	Leu	Ser	Phe	Leu	Ser	Tyr	Ser	Ile	Ala	Thr	Asn	Phe	Phe	Pro	Tyr	65	70	75	80
Ile	Ser	Ser	Pro	Leu	Ser	Tyr	Val	Ala	Trp	Leu	Val	Tyr	Trp	Leu	Phe	85	90	95	
Gln	Gly	Cys	Ile	Leu	Thr	Gly	Leu	Trp	Val	Ile	Gly	His	Glu	Cys	Gly	100	105	110	
His	His	Ala	Phe	Ser	Glu	Tyr	Gln	Leu	Ala	Asp	Asp	Ile	Val	Gly	Leu	115	120	125	
Ile	Val	His	Ser	Ala	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	130	135	140	
His	Arg	Arg	His	His	Ser	Asn	Ile	Gly	Ser	Leu	Glu	Arg	Asp	Glu	Val	145	150	155	160
Phe	Val	Pro	Lys	Ser	Lys	Ser	Lys	Ile	Ser	Trp	Tyr	Ser	Lys	Tyr	Leu	165	170	175	
Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu	Leu	180	185	190	
Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	195	200	205	
Arg	Phe	Ala	Cys	His	Tyr	Asp	Pro	Tyr	Gly	Pro	Ile	Phe	Ser	Glu	Arg	210	215	220	
Gly	Trp	Leu	Thr	Ile	Tyr	Ile	Ala	Asn	Leu	Gly	Ile	Phe	Ala	Thr	Thr	225	230	235	
Phe	Val	Leu	Tyr	Gln	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val	Met	240	245	250	255
Arg	Ile	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Cys	Phe	Leu	Val	Met	260	265	270	
Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly	Ser	275	280	285	

```

Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp
 290                               295                               300

Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val
305                               310                               315                               320

Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala
                               325                               330                               335

Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly
                               340                               345                               350

Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe
                               355                               360                               365

Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg
 370                               375                               380

Asn Lys Tyr
385

```

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (E) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: /product="synthetic oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGCTATGCG AACATGTCG AAGA

23

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (E) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..22

(D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAARTGRTGG CACFTGNGTR TC

22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2973 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(E) CLONE: pAGE2-6

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 433..520

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 521..1654

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATTCGGTAAT TCCTACATAT TTTAGAGATT AGTTTGAGTT TCCATCCAIA CTTTATTAGT

GATTATAAAT TTAAAATAGS TACTTTTCSA CTATAAAGTG AAACAAAGTA AATTAGAACG 1  
20

TGATATTAAA AAGTTAATGT TCACTGTTAT ATTTTTTTCA CAAGTAAAAA ATGGGTTATT 1  
80

TGCGGTAAAT AAAAATAACA SATATTTTGA ATTGATTAAA AAGGTTGAAA TAAGAGAGGA 2

40

GGGGAAAGAA AAGAAGGTGG GGGCCAGTA TGAAAGGGAA AGGTGTCATC AAATCATCTC 90	3
TCTCTCTCTC TACCTTCGAC CCACGGGGCG TGTCCATTTA AAGCCCTGTC TCTTGCCATT 60	3
CCCCATCTGA CCACCAGAAG AAGAGCCACA CACTCAGAAA TTAAAAAGAG AGAGAGAGAG 20	4
AGAGAGACAG AGAGAGAGAG AGATTCTGCG GAGGAGCTTC TTCTTGTAG GGTGTTTCATC 80	4
GTTATTAAACG TTATCGCCCC TACGTCAGCT CCATCTCCAG GTCCGTCGCT TCTCTTCCAT 40	5
TTCTTCTCAT TTTCGATTTT GATTCTTATT TCTTTCCAGT AGCTCCTGCT CTGTGAATTT 00	6
CTCCGCTCAC GATAGATCTG CTTATACTCC TTACATTCAA CCTTAGATCT GGTCTCGATT 60	6
CTCTGTTTCT CTGTTTTTTT CTTTGGTGG ASAATCTGAT GTTGTTTAT GTTCTGTCAC 10	7
CTTTAATAAT GATGAACTCT CTCATTGATA CAAGGATTAG TTTCTCTCGT CTACCAAACG 80	7
ATATGTTGCA TTTTCACTTT TCTTCTTTTT TTCTAAGATG ATTTGCTTTG ACCAATTTGT 40	8
TTAGATCTTT ATTTTATTTT ATTTTCTGGT GGGTTGGTGG AAATTGAAAA AAAAAAAAAA 90	9
AAAAGCATAA ATTGTTATTT GTTAATGTAT TCATTTTTTG GCTATTTGTT CTGGGTAAAA 60	9
ATCTGCTTCT ACTGTTGAAT CTTTCTGGA TTTTTACTC CTATTGGGTT TTTATAGTAA 10	10
AAATACATAA TAAAAGGAAA ACAAAGTTT TATAGATTCT CTTAAACCCC TTACGATAAA 80	10
AGTTGGAATC AAAATAATTC AGGATCAGAT GCTCTTTGAT TGATTCAGAT GCGATTACAG 10	11
TTCCATCCAA AATTTCTAG ATCGGCTGTC ACATTTTATT TTCTGTTTAA ATATCTAAAT 90	12
CTGATATATG ATGTGACAA ATTCTGGTGG CTTATACATC ACTTCAACTG TTTCTTTTTG 60	12
GCTTTGTTTG TCAACTTGGT TTTCAATACG ATTTGTGATT TCGATCGCTG AATTTTTAAT 20	13



ACAAGCAAAAC TGATGTTAAC CACAAGCAAG AGATGTGACC TGCCTTATTA ACATCGTATT 13  
 80  
 ACTTACTACT AGTCGTATTC TCAACGCAAT CTTTTTTGTA TTTCTCACAT TATGCCGCTT 14  
 40  
 CTCTACTCTT TATTCCTTTT GGTCCACGCA TTTTCTATTT GTGGCAATCC CTTTCACAAC 15  
 00  
 CTGATTTCCC ACTTTGGATC ATTTGTCTGA AGACTCTCTT GAATCGTTAC CACTTGTTTC 15  
 60  
 TTGTGCATGC TCTGTTTTTT AGAATTAATG ATAAAACTAT TCCATAGTCT TGAGTTTTCA 16  
 20  
 GCTTGTTGAT TCTTTTGCTT TTGGTTTTCT GCAGAAACAT GGSTGCAGGT GGAAGAATGC 16  
 80  
 CGGTTCTTAC TTCTTCCAAG AAATCGGAAA CCGACACCAC AAAGCGTGTG CCSTGCGAGA 17  
 40  
 AACCGCCTTT CTCGGTGGGA GATCTGAAGA AAGCAATCCC GCGGCATTGT TTCAAACGCT 18  
 00  
 CAATCCCTCG CTCTTTCTCC TACCTTATCA GTGACATCAT TATAGCCTCA TGCTTCTACT 18  
 60  
 ACGTCGCCAC CAATTACTTC TCTCTCCTCC CTCAGCCTCT CTCTTACTTG GCTTGGCCAC 19  
 20  
 TCTATTGGGC CTGTCAGGCG TGTGTCTTAA CTGGTATCTG GGTCCATAGCC CACGAATGCG 19  
 80  
 GTCACCACGC ATTCAGCGAC TACCAATGGC TGGATGACAC AGTTGGTCTT ATCTTCCATT 20  
 40  
 CTTTCTCTCT CGTCCCTTAC TTCTCCTGGA AGTATAGTCA TGGCCCTCAC CATTCCAACA 21  
 00  
 CTGGATCCCT CGAAAGAGAT GAAGTATTTG TCCCAAAGCA GAAATCAGCA ATCAAGTGGT 21  
 60  
 ACGGGAAATA CCTCAACAAC CCTCTTGGAC GCATCATSAT GTTAACCGTC CAGTTTGTCC 22  
 20  
 GCGGGTGGCC CTTGTACTTA GCGTTTAAAG TCTCTGSCAS ACGTATCAC GGGTTGGCTT 22  
 80  
 GGCATTTCTT CCGCAAGGCT CCGATCTACA ATGACCGAGA ACGCTCCAS ATATACTCT 23  
 40  
 TTGATGCGGG TATTCTAGCC GTCTSTTTTG STTTTADCC TTACGCTGCT GCACAAGGGA 24  
 00  
 TGGCCTCGAT GATCTGCCTC TACGAGTAC CGCTTCTGAT AGTGAATGCG TTCTCTGTCT 24  
 60

```

TGATCACTTA CTTGCAGCAC ACTCATCCCT CGTTGCCTCA CTACGATTCA TCAGAGTGGG 25
30
ACTGGCTCAG GGGAGCTTTG GCTACCGTAG ACAGAGACTA CGGAATCTTG AACAAAGGTGT 25
36
TCCACAACAT TACAGACACA CACGTGGGTC ATCACTGTG CTCGACAATG CCGCATTATA 26
40
ACGCAATGGA AGCTACAAAG GCGATAAAGC CAATTCTGGG AGACTATTAC CAGTTCGATG 27
00
GAACACCGTG GTATGTGGCG ATGTATAGGG AGGCAAAGGA GTGTATCTAT GTAGAACCGG 27
60
ACAGGGAAGG TGACAAGAAA GGTGTGTA CT GGTACAACAA TAAGTTATGA GGATGATGGT 28
20
GAAGAAATTG TCGACTTTTC TCTTGTCTGT TTGTCTTTTG TTAAAGAAGC TATGCTTCGT 28
80
TTTAATAATC TTATTGTCCA TTTTGTGTG TTATGACATT TTGGCTGCTC ATTATGTTAT 29
40
GTGGGAAGTT AGCGTTCAAA TGTTTTGGGT CGG 29

```

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME KEY: misc feature
  - (B) LOCATION: 1..23
  - (C) OTHER INFORMATION: /product=
- "synthetic  
oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGCATGTNG ARAANAARTG FTR 23

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..23

(D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGGCATGTRC TRAANAERTG RTG

23